



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 15/63, 1/21, 5/00, C07K 7/00, 14/435		A1	(11) International Publication Number: WO 99/66041
			(43) International Publication Date: 23 December 1999 (23.12.99)
(21) International Application Number: PCT/US99/13418		20851 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, MD 22020 (US). BREWER, Laurie, A. [US/US]; Apartment 115, 410 Van Dyke Street, St. Paul, MN 55119-4321 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). MUCENSKI, Michael [US/US]; 3263 Mandale Drive, Cincinnati, OH 45239 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace, #316, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place, #24, Gaithersburg, MD 20878 (US). SHI, Yangu [CN/US]; 437 West Side Drive, #102, Gaithersburg, MD 20878 (US). MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). KOMATSOULIS, George [US/US]; 9518 Garwood Street, Silver Spring, MD 20901 (US).	
(22) International Filing Date: 15 June 1999 (15.06.99)			
(30) Priority Data:			
60/089,509 16 June 1998 (16.06.98) US 60/089,510 16 June 1998 (16.06.98) US 60/089,508 16 June 1998 (16.06.98) US 60/089,507 16 June 1998 (16.06.98) US 60/090,112 22 June 1998 (22.06.98) US 60/090,113 22 June 1998 (22.06.98) US			
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).			
(72) Inventors; and		(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).	
(75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). WEI, Ying-Fei [CN/US]; 242 Gravett Drive, Berkeley, CA 94705 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD			
		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>	
(54) Title: 94 HUMAN SECRETED PROTEINS			
(57) Abstract			
<p>The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

94 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of

the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

5

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and
10 polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

15

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original
20 environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

25

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing
30 to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of the coding sequence, but do not comprise all or a portion of any intron.

- 5 In another embodiment, the nucleic acid comprising the coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene in the genome).

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone
10 deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a
15 molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID
20 NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of
25 microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an
30 overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's

solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions.

- 5 Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M
- 10 NaH_2PO_4 ; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

- Note that variations in the above conditions may be accomplished through the
- 15 inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above,
- 20 due to problems with compatibility.

- Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid
- 25 molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

- The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of
- 30 single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA

that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA
5 backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids
10 joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more
15 detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of
20 modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a
25 heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation,
30 iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

(See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

Preferred polypeptides of the invention comprise the following amino acid sequence: TRPEKVQAPLKWFKFQILDPP (SEQ ID NO:249). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in dendritic cells and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, nervous system, and inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing

immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in dendritic cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, and autism. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Furthermore, expression of this gene product in primary dendritic cells also indicates that it may play a role in mediating responses to infection and controlling immunological responses, such as those that occur during immune surveillance. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 885 of SEQ ID NO:11, b is an integer of 15 to 899, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of this gene share homology with the Tbc1 gene of Mus musculus which is thought to play a role in the cell cycle and differentiation of various tissues (See Genebank accession no. gi|988221 as well as Medline article
 5 no.96032578; all references available through these accessions are hereby incorporated by reference herein). One embodiment for this gene is the polypeptide fragments comprising the following amino acid sequence:

SAEFGVAPLPGRRGSPVRQLAQFRRRLLRGSGGRGAPGRPPRCPGEARVMXPPSCIQDEPFPHPLEPEP
 GVSAQPGPGKPSDKRFRLLWYVGGSCLDHRTTLPMLPWLMAEIRRRSQKPEAGGCGAPAAREVILVLSAP
 10 FLRCVPAPGAGASGGTSPSATQPNPAVFIFEHKAQHISRFIHNSHDLTYFAYLIKAQPDDPESQMACHV
 FRATDPSQVPDVISSIRQLSKXAMKEDAKPSKDNEAFYNSQKFEVLYCGKVTVTPQEGPLKPHR
 (SEQ ID NO:250); PMLPWLMAEIRRRS (SEQ ID NO:251); IHNSHDLTYFAYLIKAQPD
 (SEQ ID NO:252); KFEVLYCGKVTV (SEQ ID NO:253); and/or ISSIRQLSKAMKE
 (SEQ ID NO:254). Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed primarily in smooth muscle and dendritic cells and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
 20 not limited to, cardiovascular diseases and immune and inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and cardiovascular system, expression of this gene at significantly
 25 higher or lower levels is routinely detected in certain tissues or cell types (e.g., smooth muscle and dendritic cells, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from
 30 an individual not having the disorder.

The tissue distribution in smooth muscle and dendritic cells and homology to a protein involved in regulation of cell cycle and tissue differentiation indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

detection/treatment and/or prevention of immune system disorders, cardiovascular disorders or diseases, including cancer and other proliferative disorders. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders.

- 5 Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation
10 of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such
15 as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus
20 erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of
25 various blood lineages, and in the differentiation and/or proliferation of various cell types.

- Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or
30 embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and

such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1126 of SEQ ID NO:12, b is an integer of 15 to 1140, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of this gene shares sequence homology with alpha-1 antitrypsin (See Genebank accession no. gn|PID|d1021080; all references available through this accession are hereby incorporated by reference herein). Alpha-1-antitrypsin is an important plasma protease inhibitor affecting a wide variety of serine proteases involved in coagulation, fibrinolysis and kinen generation.

Preferred polypeptides of the invention comprise the following amino acid sequence: GERRNWGGEVYYSTGYSSRK (SEQ ID NO:255). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in healing groin wound and to a lesser extent in some other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, wound healing disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the healing groin wound, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., healing, regenerative, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 132 as residues: Phe-25 to Tyr-30, Gln-37 to Arg-42, Lys-106 to Leu-112, Leu-123 to Leu-130, Gln-142 to Phe-150, Gln-183 to Lys-188, Asp-219 to Glu-226, Lys-359 to Glu-366. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in healing groin wound and homology to alpha-1 antitrypsin indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and therapeutic treatment of wound healing disorders. In addition, since healing wounds have transcriptional environments similar to developing tissues, The translation product of this gene is useful for the diagnosis and treatment of cancer and other proliferative disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1431 of SEQ ID NO:13, b is an integer of 15 to 1445, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The translation product of this gene shares homology with members of the HEMK family of modification methylases (See, e.g., Genbank Accession No. gb|AAD26417.1|AF131220_1; all references available through this accession are

10 hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: EPGAAQESW (SEQ ID NO:256); LCARPSCSYTGAENQGQPRSPGWGSSHVWGWWG VGSPPFLGSQEWGLAPDLPDQEEQPVGRHSCPDMSQCIKRGHPVGFSKHAWRCLVGCCPWEEEKRSCHPF
 15 GAXLLWVLRFAALQXPXYEDPAALDGGEEGMDIXTHILALAPRLKDSGSIFLEVDPRHPXLVSSWL QSRPDLYLNLVAVRRDFCGRPRFLHIRRS GP (SEQ ID NO:257); LCARPSCSYTGAENQGQPR SPGWGSSHVWGWWGVGSP (SEQ ID NO:258); FLGSQEWGLAPDLPDQEEQPVGRHSCPDMS QCIKR (SEQ ID NO:259); GHQPVGFSGHAWRCLVGCCPWEEEKRSCHPF
 20 GAXLLW (SEQ ID NO:260); VLRFAALQXPXYEDPAALDGGEEGMDIXTHILALAPRL (SEQ ID NO:261); and/or LKDSGSIFLEVDPRHPXLVSSWLQSRPDLYLNLVAVRRDFCGRPRFLHIRRS GP (SEQ ID NO:262). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in immune and tumor tissues, and to a lesser extent in some other tissues such as heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
 25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and inflammatory disorders and tumorigenesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and
 30 tumor tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative

to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 133 as residues: Met-1 to Cys-6, Ser-26 to Gly-35.

5 Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in tumors of immune origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of such tumors, in addition to other tumors where expression has been indicated. Additionally, this gene is a good target for antagonists, particularly
10 small molecules or antibodies, which block binding of the receptor by its cognate ligand(s). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
15 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1194 of SEQ ID NO:14, b is an integer of 15 to 1208, where both a and b correspond to the positions of nucleotide
25 residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

The translation product of this gene shares sequence homology with mouse von Ebner minor salivary gland protein which may play a role in carbohydrate
30 metabolism (See Genebank Accession No. gb|AAA87581.1|; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: QELLVKIPLDMVAGFNTPL (SEQ ID NO:263); LRIQLLHKLSFLVNALAK QVMNLLVP (SEQ ID NO:264); AGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEK LTQELKDHNATSILQQLPLL (SEQ ID NO:266); and/or HXIWLKVITXNILQLQVKPS (SEQ ID NO:265). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in respiratory tissues such as trachea, larynx and other pulmonary tissues, and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, respiratory system and oral disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 134 as residues: Lys-39 to Asn-48, Arg-63 to Gly-68, Pro-101 to Gln-106. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution combined with the homology to von Ebner minor salivary gland protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of respiratory and oral diseases. Furthermore, The tissue distribution in pulmonary tissues also indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of tumors within these tissues, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the

above listed tumors and tissues. Protein may show utility in the diagnosis, treatment, and/or prevention of disorders in carbohydrate metabolism.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1161 of SEQ ID NO:15, b is an integer of 15 to 1175, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The gene encoding the disclosed cDNA is believed to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2.

This gene is expressed primarily in fast-growing tissues such as fetal tissues, hematopoietic cells and tumor tissues and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, growth disorders, tumorigenesis, and immune or inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fast-growing tissues such as fetal tissues, hematopoietic cells and tumor tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a

disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fast growing tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression in embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation or cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages which implicates the protein product of this gene as being useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. Thus, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2360 of SEQ ID NO:16, b is an integer of 15 to 2374, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

The translation product of this gene shares sequence homology with mitochondrial NADH-Ubiquinone oxidoreductase, chain 2.

5 Preferred polypeptides of the invention comprise the following amino acid sequence: HFIITLTTFFTNYFL (SEQ ID NO:267); and/or MKITFQDLFPMWNSFKCFL HGNVPSLFLVLFPLLTCFSFPYTVNSGTKLDWVGWLVGWFFLEFMYINKGFEVTSENNISKRVLVRENIR IKSSPERVLRM (SEQ ID NO:268). Polynucleotides encoding these polypeptides are also provided.

10 This gene is expressed primarily in stromal cells (cell code TF274), induced epithelial cells and human cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic disorders and conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, brain, and integument, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in epithelial and cerebral tissues combined with the homology to a known mitochondrial NADH-Ubiquinone oxidoreductase gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sach's disease, phenylketonuria, galactosemia, porphyrias, and Hurler's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

30

supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1581 of SEQ ID NO:17, b is an integer of 15 to 1595, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

The translation product of this gene shares sequence homology with Platelet activating factor acetylhydrolase which inactivates Platelet activating factor, a potent phospholipid mediator affecting various physiological processes (See, e.g., Genbank Accession Nos. gi|349824|gb|AAA02880.1| and gi|2072303|gb|AAC04610.1|; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: RFWGSYEPHFSQEVSVIPP (SEQ ID NO:269); and/or IRGNYFSGRKKSSSDT PKGSKDKISVWNRSQXACIRICKVHPNYIQIYLWHSATSF (SEQ ID NO:270). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in CD34 depleted buffy coat (cord blood) and to a lesser extent in human prostate cancer, stage 3 fraction.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, particularly of the prostate. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of

disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., prostate, cancerous and wounded tissues) or bodily fluids (e.g., lymph, cord blood, serum, plasma, urine, synovial fluid and spinal fluid) or
5 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in CD34 depleted buffy coat combined with the homology to Platelet-activating factor acetylhydrolases, proteins involved in
10 regulation of platelet activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in hematopoietic cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells.
15 Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g. by boosting immune responses.

20 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or
25 receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1273 of SEQ ID NO:18, b is an integer of 15 to 1287, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

Preferred polypeptides of the invention comprise the following amino acid sequence: AGNQVEPFHVSLPSCLSPLPHLGHSMGVPSPTAWPSLASFHTQKKARIRQEEES PPLPSPQELAFSALRVFFRV (SEQ ID NO:271). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in primary dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunosuppression and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 138 as residues: Arg-20 to Lys-44, Arg-59 to Arg-68, Trp-74 to Lys-86, Thr-91 to Val-102. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in primary dendritic cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment

of a variety of immune system disorders. Expression of this gene product in dendritic cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere
5 herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the
10 treatment of cancer e.g. by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Furthermore, the protein may also be used to determine
15 biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
25 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1382 of SEQ ID NO:19, b is an integer of 15 to 1396, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

The translation product of this gene shares sequence homology with peptide/histidine transporter from *Rattus norvegicus* and other peptide transporters which are thought to be important in transporting amino acids and peptides into cells (See, e.g., Genbank Accession No. gb|AAD24570.1|AF121080_1; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: FIQQNISFLLGYSIPVGCVGLAFFIFLFATPVFITKPP (SEQ ID NO: 272).

Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

This gene is expressed primarily in macrophages and to a lesser extent in other immune cells including primary dendritic cells, neutrophils, resting T-cells, B cell lymphomas) and lung and fetal liver spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and disorders, particularly of the immune and hematopoietic systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 139 as residues: Arg-23 to Gln-30, Asp-37 to Asp-50, Glu-230 to Met-235, Pro-271 to Arg-281, Arg-306 to Ser-316, Ser-318 to Gly-325. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in macrophages and other immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses). Alternatively expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation or cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1263 of SEQ ID NO:20, b is an

integer of 15 to 1277, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

5 The translation product of this gene shares sequence homology with procollagen-proline dioxygenase, an apparently secreted protein which is thought to be important in the formation of 4-hydroxyproline in collagens (See, e.g., Genbank Accession No. pir|A33832|DACHA; all references available through this accession are hereby incorporated by reference herein). Furthermore, the translation product has
10 an EF-hand domain (Prosite PS00018) which is a calcium binding domain as found in calmodulin, calpain, spectrin alpha chain, etc., (See, e.g. GeneSeq Accession No.R78523; all references available through this accession are hereby incorporated by reference herein).

 Preferred polypeptides of the invention comprise the following amino acid
15 sequence:
VSAHHPGADGVTAXQILPTEEYEEAMSTMQVSQDLFRLLDQNRDGHQLQREVLAQTRLGNGWWMTP
ESIQEMYAAIKADPDGDGVLSQLQEFNSMDLRDFHKYMRSHKAESSELVRNSHHTWLYQGEGAHHIMRAI
RQVRRLRLTRLSPEIVELSEPLQVVRYGEGGHYHAHVDSGPVYPETICSHTKLVANESVVPFETSCRYMTV
LFYLNNTGGGETVFPVADNRTYDEMSLIQDDVDLRDTRRHCDKGNLRVKPQQGTAVFWYNYLPDGQGW
20 VGDVDDYSLHGGCLVTRGTKWIANNWINVDPSRARQALFQQEMARLAREGGTDSQPEWALDRAXXDARV
EL (SEQ ID NO:273); AVFWYN (SEQ ID NO:274); TVLFYLNNTGGGETVFP (SEQ
ID NO:275); DLFRLLDQNRDGHQLQREVLAQTRLGNGWWMTPESIQEMYAAIKADPDGDGVL
LQEFNS (SEQ ID NO:276); VSAHHPGADGVTAXQILPTEEYEEAMSTMQVSQDL (SEQ ID
NO:277), FRLLDQNRDGHQLQREVLAQTRLGNGWWMTPESIQEMY (SEQ ID NO:278);
25 AAIKADPDGDGVLSQLQEFNSMDLRDFHKYMRSHKAESS (SEQ ID NO:279); ELVRNSHHTWLY
QGEGAHHIMRAIRQVRRLRLTRLSPEI (SEQ ID NO:280); VELSEPLQVVRYGEGGHYHAHVDS
GPVYPETICSHTKL (SEQ ID NO:281); VANESVVPFETSCRYMTVLFYLNNTGGGETVFPV
DNR (SEQ ID NO:282); TYDEMSLIQDDVDLRDTRRHCDKGNLRVKPQQGTAVFW (SEQ ID
NO:283); YNYLPDGQGWVGDVDDYSLHGGCLVTRGTKWIANNWIN (SEQ ID NO:284);
30 and/or VDPSSRARQALFQQEMARLAREGGTDSQPEWALDRAXXDARVEL (SEQ ID NO:285).

Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in human endometrial tumor and to a lesser extent in brain, as well as a variety of other normal and cancerous tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometrial cancer, in addition to other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and neural systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, reproductive, and/or other tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid, lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 140 as residues: Ser-21 to His-33, Ala-35 to Thr-43. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in endometrial tumors combined with the homology to procollagen-proline dioxygenase indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment and prevention of these tumors, in addition to other tumors where expression has been indicated. The polypeptides of the invention is a good target for antagonists, particularly small molecules or antibodies, which block binding of the receptor by its cognate ligand(s). Accordingly, preferred are antibodies and or small molecules which specifically bind an extracellular portion of The translation product of this gene. Also provided is a kit for detecting endometrial cancer. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting endometrial cancer in an individual which comprises a step of contacting an antibody specific for The translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining

whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. Additionally, the homology to a conserved collagen metabolizing protein would suggest that this protein may also be important in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1767 of SEQ ID NO:21, b is an integer of 15 to 1781, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 12

This gene is expressed primarily in human osteoblastoma cell lines (5/23 unique sequences) and to a lesser extent in T cells (4/23).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, osteoblastoma, and other bone-related disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing

immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., bone and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tumors of bone origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Additionally, this gene is a good target for antagonists, particularly small molecules or antibodies, which block binding of the receptor by its cognate ligand(s). Accordingly, preferred are antibodies and or small molecules which specifically bind an extracellular portion of The translation product of this gene. The extracellular regions can be ascertained from the information regarding the transmembrane domains as set out above. Also provided is a kit for detecting osteoblastoma and other bone related cancers. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting bone related cancers in an individual which comprises a step of contacting an antibody specific for The translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
5 formula of a-b, where a is any integer between 1 to 1477 of SEQ ID NO:22, b is an integer of 15 to 1491, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

10 The translation product of this gene is a human homolog of the mouse acetylcholine receptor gamma chain, and is almost identical to a human acetylcholine receptor gamma chain (See, e.g., Genbank Accession Nos.: emb|CAA27442.1| and gb|AAA51568.1|; all references available through these accessions are hereby incorporated by reference herein) which is thought to be important in transmission of
15 nerve impulses to muscles.

Preferred polypeptides of the invention comprise the following amino acid sequence: LLADLMRNYDPLRP (SEQ ID NO:286); ISVTYFPFDWQNC SLIFQS (SEQ ID NO:287); SMARGVRKVFLRLLPQ (SEQ ID NO:288); QASPAIQACVDACNL MAR (SEQ ID NO:289); and/or YNQVPDLPFPGDPRPYL (SEQ ID NO:290). Polynucleotides
20 encoding these polypeptides are also provided. This gene maps to chromosome 2, and therefore, is used as a marker in linkage analysis for chromosome 2. Included in this invention as preferred domains are Neurotransmitter-gated ion-channels domains, which were identified using the ProSite analysis tool. Structurally, members of the family of Neurotransmitter-gated ion-channels are composed of a large extracellular
25 glycosylated N-terminal ligand-binding domain, followed by three hydrophobic transmembrane regions which form the ionic channel, followed by an intracellular region of variable length. A fourth hydrophobic region is found at the C-terminal of the sequence. In the N-terminal extracellular domain of AchR/GABA/5HT3/Gly receptors, there are two conserved cysteine residues, which, in AchR, have been
30 shown to form a disulfide bond essential to the tertiary structure of the receptor. A number of amino acids between the two disulfide-bonded cysteines are also conserved. We have therefore used this region as a signature pattern for this subclass

of proteins. The consensus pattern is as follows: C-x-[LIVMFQ]-x-[LIVMF]-x(2)-[FY]-P-x-D-x(3)-C.

Preferred polypeptides of the invention comprise the following amino acid sequence: CSISVTYFPFDWQNC (SEQ ID NO:291). Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides comprising the Neurotransmitter-gated ion-channel domain of the amino acid sequence referenced in Table 1 for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of the amino acid sequence referenced in Table 1 for this gene. The additional contiguous amino acid residues is N-terminal or C-terminal to the Neurotransmitter-gated ion-channel domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and C-terminal to the Neurotransmitter-gated ion-channel domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to Neurotransmitter-gated ion-channels.

This gene is expressed primarily in fetal tissues (56/58 unique sequences), specifically lung (42/58) and Dura Mater (14/58). It was also detected (1 sequence each) in a differentially expressed human cerebellum library and human tonsil library

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly fetal lung and brain, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues and cell types (e.g., developmental, neural, differentiating, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid, pulmonary surfactant) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 142 as residues: Met-1 to Pro-7, Gln-21 to Glu-27,

Arg-35 to Asp-49, Asn-66 to Leu-72, Trp-82 to Glu-95, Pro-158 to Asn-163.

Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in dura mater combined with the homology to a conserved acetylcholine receptor indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, and/or disorders of the cardiovascular and pulmonary systems. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1825 of SEQ ID NO:23, b is an integer of 15 to 1839, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

Preferred polypeptides of the invention comprise the following amino acid sequence: VLKYALFLVLKNYYYCPY (SEQ ID NO:292). Polynucleotides

5 encoding these polypeptides are also provided.

This gene is expressed primarily in small intestine and to a lesser extent in lung cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal and pulmonary disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the intestinal and pulmonary
15 systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., gastrointestinal, pulmonary, and/or other tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, lymph, and/or pulmonary surfactant) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,
20 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in small intestine indicates a role in the detection and/or treatment of gastro-intestinal disorders including Whipple's disease, Ulcers, and indigestion. Expression in the lung indicates a potential role in the treatment and/or
25 detection of certain pulmonary defects such as pulmonary edema and embolism, bronchitis, cystic fibrosis and lung cancer. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed
30 against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1370 of SEQ ID NO:24, b is an integer of 15 to 1384, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

In another embodiment, polypeptides of the invention comprise the following amino acid sequence:

MREYGVERDLAVYNQLLNIFPKEVFRPRNIIQRIFVHYPRQCEGIAVLEQMENHGVM PNKETEFLLIQ
IFGRKSY PMLKLVRLKLWFFRPMNVNPFVPRDLPQDPVELAMFGLRHMEPDLSARVTIYQVLPKDST
GAADPPQPHIVGIQSPDQQAALARHNPARPVFVEGPFSLWLRNKC VYYHILRADLLPPEEREVEETPEE
WNLYYPMQLDLEYVRSGWDNYEFDINEVEEGPVFAMCMAGAHDQATMAKWIQGLQETNPTLAQIPVVFR
LAGSTRELQTSSAGLEEPPLPEDHQEEDNLQRQQGQS (SEQ ID NO:293).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain and to a lesser extent in pancreas, testes, and other tissue types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, behavioral, gastrointestinal, and endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., brain, endocrine, and/or other tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, and lymph) or another tissue or cell sample taken from an individual having such a

disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 144 as residues: Val-33 to Arg-39, Ser-57 to Thr-66,
5 Pro-80 to Lys-86, Pro-155 to Cys-160, Val-215 to Pro-223, Pro-250 to Gly-255, Pro-311 to Glu-323, Arg-338 to Tyr-344, Ser-396 to Gln-401, Pro-410 to Ser-431.

Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative
10 disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment
15 and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
20 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of
25 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1667 of SEQ ID NO:25, b is an
30 integer of 15 to 1681, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of this gene shares sequence homology with the acid labile subunit of the insulin like growth factor binding subunit which is thought to be important in modulating the activity of Insulin like growth factor. In addition, this
5 gene also shares homology with the melibiose carrier protein (thiomethylgalactoside permease II) of *Caenorhabditis elegans* (See Genebank Accession No. gi|1280135; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid
10 sequence: FQFGWASTQISHLSLIPEL (SEQ ID NO:294); LRYAFTVVANITVY (SEQ ID NO:295); FVYGSMFLDKVANGLA (SEQ ID NO:296); WHLVGTVCVLLSFFPFIF (SEQ ID NO:297); and/or GHFLNDLCASMWFTY (SEQ ID NO:298). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in macrophages and to a lesser extent in
15 dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic disorders. Similarly, polypeptides and
20 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and/or immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g.hematopoietic, immune, and/or
25 other tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, and lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
30 epitopes shown in SEQ ID NO: 145 as residues: Ala-28 to Ala-33, Arg-38 to Leu-48, Thr-120 to Lys-125, Gly-155 to Gln-163, Gly-200 to Glu-214. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution predominantly in dendritic cells and macrophages combined with homology to a growth factor binding subunit indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1935 of SEQ ID NO:26, b is an integer of 15 to 1949, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

The translation product of this gene was shown to have homology to the T13C5.6 gene product from *Caenorhabditis elegans* (See Genebank Accession No. gi|1049369; all references available through this accession are hereby incorporated by reference herein).

5 Preferred polypeptides of the invention comprise the following amino acid sequence: AIPLRVLVVLWAFVLGLSRVMLGRHNVTDVAFGFFLGVMQ (SEQ ID NO:299); and/or VGLSRVLGRHTDV (SEQ ID NO:300). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in placenta and small intestine.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pregnancy, reproductive, and/or gastrointestinal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
15 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the intestinal and endocrine systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, gastrointestinal, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid
20 and spinal fluid, amniotic fluid,) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in placenta indicates a potential role for this protein in
25 the detection and/or treatment of pregnancy disorders such as miscarriage and/or gastro-intestinal disorders such as indigestion, ulcers and Whipple's disease. Alternatively, polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylketonuria, galactosemia, porphyrias, and Hurler's
30 syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2272 of SEQ ID NO:27, b is an integer of 15 to 2286, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

Preferred polypeptides of the invention comprise the following amino acid sequence: SFYKMKRNSYDRLRKVV (SEQ ID NO:301). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in prostate and spleen and to a lesser extent in most cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, immune, and/or other tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, seminal fluid,

and lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution in prostate indicates a potential role in the treatment and/or detection of prostate disorders including benign prostate hyperplasia and prostate cancer. Expression in spleen indicates a role in the treatment and/or detection of spleen disorders such as splenitis and spleen cancer. Alternatively, the expression in the spleen may suggest that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system
10 disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved
15 in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g. by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for
20 immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may
25 show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Many polynucleotide sequences, such as EST sequences, are publicly
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 516 of SEQ ID NO:28, b is an integer of 15 to 530, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene was shown to have homology to both a human IgE-binding protein as well as to the human gene for Human Factor XIII (See Genebank Accession Nos. gb|S76337|S76337 and Q25893, respectively).

Preferred polypeptides of the invention comprise the following amino acid sequence: LHQLRPPHRFPLIPPAAGAGAPPGCGYCVFWLLNPLP (SEQ ID NO:302), and/or MPWKRAVLLMLWFIGQAMWLAPAYVLEFQGKNTFLFIWLAGLFLLINCSILIQIISH YKEEPLTERIKYD (SEQ ID NO:303). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and behavioural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, immune, and/or other tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, and lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in infant brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and

5 "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism,

10 and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

15 immunotherapy targets for the above listed tumors and tissues. Alternatively, considering the homology to a conserved human gene for IgE as well as to a conserved blood clotting factor may suggest this gene is useful for the diagnosis and treatment of a variety of immune system disorders. Homology of this gene to a blood clotting factor, specifically, indicates a role in the regulation of the proliferation;

20 survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g., by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene

25 product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a

30 nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. In addition, this gene product may have commercial utility in the expansion of

stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1282 of SEQ ID NO:29, b is an integer of 15 to 1296, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

Preferred polypeptides of the invention comprise the following amino acid sequence: ARAQFFAFQLRPAPGRPGSPVA (SEQ ID NO:304);
AGLPGALTAPAXHHHADSRAELVVQPLSPPRLLSHAGLASAAGASSLXRVPGEAESLCALSPGSALR
FPAASCSRXPXREPSGDEGTAGALPSPWLAALGPGGRPAVRRVLPRLGGRAGQLPRGLPVPRGLRHAGRY
HLLRLLRAPLLLRGRRQAGAGRLHQRPPRTGAPRHCAACLRPLSHRRLHLHCVHHPGLCSGYLLHL
FETQGALAAANPLLTPQLSDRDPADHPDLHQPGQTLPAVQHSHELQLHRLHPQVLLSHLVSWCHPSI
SLTPFSRSPHWLGRAVQTFSSX (SEQ ID NO:305); AGLPGALTAPAXHHHADSRAELVVQP
LSPPRLLSHA (SEQ ID NO:306); GLASAAGASSLXRVPGEAESLCALSPGSALRFPAASCSR
(SEQ ID NO:307); XREPSGDEGTAGALPSPWLAALGPGGRPAVRRVLPRLGGR (SEQ ID
NO:308); AGQLPRGLPVPRGLRHAGRYHLLRLLRAPLLLRGRRQAG (SEQ ID NO:309);
AGRLHQRPPRTGAPRHCAACLRPLSHRRLHLHCVHHPGL (SEQ ID NO:310); CSGYLLHLF
ETQGALAAANPLLTPQLSDRDPADHPDLHQ (SEQ ID NO:311); and/or PQGTLPVQHS
ELQLHRLHPQVLLSHLVSWCHPSISLTPFSRSPHWLGRAVQTFSSX (SEQ ID NO:312).

Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed primarily in heart and to a lesser extent in the embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and developmental systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cardiopulmonary, developmental, and/or other tissues) or bodily fluids (e.g., lymph, sputum, serum, plasma, urine, synovial fluid and spinal fluid, amniotic fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 149 as residues: Gln-23 to Gly-30, Gln-35 to Gln-43, Leu-73 to Glu-84, Arg-125 to Pro-133, Ser-140 to Thr-145, Thr-153 to Thr-164. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in heart indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or detection of a range of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, embolism, vasculitis, myocardial infarction, myocarditis, ischemia, stroke, in addition to developmental and metabolic disorders. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Alternatively, the expression in embryonic tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Furthermore, protein may play a role in the regulation of cellular division. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early

hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Furthermore, the protein
5 may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
15 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1965 of SEQ ID NO:30, b is an integer of 15 to 1979, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in human teratocarcinoma cell line treated with retinoic acid and human brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental abnormalities and neural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
30 a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, differentiating, neural, and/or other

tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid, amniotic fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution in teratocarcinoma cell line indicates that polynucleotides and polypeptides corresponding to this gene are useful for early diagnosis and treatment of developmental abnormalities, including agenesis, aplasia, hypoplasia, dysraphic anomalies, division failures, dysplasia, etc. Additionally, the gene and its expression can be used for teratogen detection or classification.
- 10 Alternatively, considering the expression within human brain tissue may suggest that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic
- 15 disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
- 20 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly
- 25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
- 30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1260 of SEQ ID NO:31, b is an

integer of 15 to 1274, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

5 The translation product of this gene was shown to have homology to the human B-cell growth factor which is known to be involved in the maturation of B-cells (See Genbank Accession No. gi|522145; all references available through this accession are hereby incorporated by reference herein).

10 Preferred polypeptides of the invention comprise the following amino acid sequence: VAHTCNLSTLGGQGGRIERTAGQEFKTS (SEQ ID NO:313).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in multiple sclerosis and prostate tissues and to a lesser extent in brain and osteoblasts.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, muscle, reproductive, and neural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of
20 disorders of the above tissues or cells, particularly of the central nervous system and/or PNS, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., muscle, reproductive, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid, seminal fluid) or another tissue or cell sample taken from an individual having
25 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 151 as residues: Gln-28 to Asp-35. Polynucleotides encoding said polypeptides are also provided.

30 The tissue distribution in multiple sclerosis indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory

conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1517 of SEQ ID NO:32, b is an integer of 15 to 1531, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

The translation product of this gene was shown to have homology to the B0035.14 gene of *Caenorhabditis elegans* (See, e.g., Genbank Accession No. gnl|PID|e242592; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: TIKMQTENLGVVYYVNKDF (SEQ ID NO:314); MVSNNPPY (SEQ ID NO:316); HASEL (SEQ ID NO:317); and/or VEEDYVTNIRNNC (SEQ ID NO:315). Polynucleotides encoding these polypeptides are also provided.

5 This gene is expressed primarily in bone marrow and to a lesser extent in lung and various tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic, and/or cardiopulmonary disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., proliferating, haematopoietic, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid, pulmonary surfactant) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 152 as residues: Ile-34 to Glu-39, Lys-49 to Lys-56, Val-63 to Glu-68, Thr-73 to Asp-88, Arg-97 to Pro-107. Polynucleotides encoding said polypeptides are also provided.

25 The tissue distribution in bone marrow indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or

chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency, etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and
5 in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
10 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2076 of SEQ ID NO:33, b is an integer of 15 to 2090, where both a and b correspond to the positions of nucleotide
20 residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

Preferred polypeptides of the invention comprise the following amino acid sequence: LVALDRMEYVRTFRKREDLRGRLFWALDILLDLLD (SEQ ID NO:318).
25 Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in T-cells and breast cancer tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
30 not limited to, immune disorders and breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of

disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, breast, proliferating, and/or other tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, breast milk, and lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 153 as residues: Tyr-105 to Pro-113, Gln-122 to Pro-133, Pro-140 to Asp-155. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in T cells and breast cancer indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g., by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the

differentiation and/or proliferation of various cell types. The expression of the gene in the breast cancer tissue may indicate T-cell mediated immune reaction to the cancer tissue.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 992 of SEQ ID NO:34, b is an integer of 15 to 1006, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 25

The translation product of this gene shares sequence homology with an yeast ankyrin repeat-containing protein Akr1p which is thought to be important in pheromone response pathway (See Genebank Accession No. gi|466522; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: SVALFYNFGKSWKSDPGIIKXTEEQKKKTIVELAEATGSLDLSIFCSTCLIRKPVRSK HCGVCNRCIAKFDHHCWPWGNVCV GAGNHRYP (SEQ ID NO:319); FDHHCWPWGNVCV (SEQ ID NO:320); and/or QMYQISCLGITTNERNARR (SEQ ID NO:321). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

This gene is expressed primarily in human lung cancer cells, B-cell lymphoma and to a lesser extent in fetal tissues and tumor cells of various origins.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, cancer of various origins, particularly of the lungs and hematopoietic systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lung, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., lung, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, pulmonary surfactant, and lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 154 as residues: Thr-28 to Phe-35, Asp-140 to Ser-145. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in lung cancer indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in lymphomas indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g., by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, distribution

in tumor tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancers of various origins, especially lung B-cell lymphoma, stomach cancer, osteoclastoma. Additionally, this gene is a good target for antagonists, particularly small molecules or antibodies, which block binding of the receptor by its cognate ligand(s). Accordingly, preferred are antibodies and or small molecules which specifically bind an extracellular portion of The translation product of this gene. Also provided is a kit for detecting lung cancer. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting lung cancer in an individual which comprises a step of contacting an antibody specific for The translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. Furthermore, the protein may also be used to determine biological activity; raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1773 of SEQ ID NO:35, b is an integer of 15 to 1787, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 26

The gene encoding the disclosed cDNA is believed to reside on chromosome 15. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 15.

This gene is expressed primarily in infant brain and to a lesser extent in a variety of other tissues and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and neurodegenerative diseases of the brain and nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, CNS, and/or PNS, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developmental, differentiating, neural, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 155 as residues: Ser-33 to Ile-41. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in infant brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia,

mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in

5 normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a

10 nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

15 related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

20 formula of a-b, where a is any integer between 1 to 1187 of SEQ ID NO:36, b is an integer of 15 to 1201, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

25 The translation product of this gene shares sequence homology with a zinc transporter, ZnT-1, which is thought to regulate zinc excretion from cells and maintain homeostasis (See Genebank Accession No. gb|AAA79234.1|, all references available through this accession are hereby incorporated by reference herein; as well as Palmiter and Findley, EMBO J. 14:639-649 (1995), which is hereby incorporated

30 by reference herein). Transformation of normal cells with a mutant rat ZnT-1 lacking the first membrane-spanning domain conferred zinc sensitivity on wild-type cells, suggesting that ZnT-1 functions as a multimer. Deletion of the first two membrane-

spanning domains resulted in a non-functional molecule, whereas deletion of the C-terminal tail produced a toxic phenotype. Transmembrane domains of the protein of the current invention are predicted using PSORT to comprise the following amino acid residues of the amino acid sequence referenced in Table 1 for this gene: Ser-42 to Ala-58, Ala-83 to Leu-99, Leu-115 to Gly-131, Val-249 to Val-265, and/or Val-314 to Leu-330. Therefore, preferred polypeptides of the present invention are the predicted extracellular domains, comprising the following amino acid sequence:

RVTSSLAMLSDS (SEQ ID NO:322); AIERFIEPHEMQQPL (SEQ ID NO:323); and/or NALVFYFSWKGCSEGDFCVNPCFPDPCKPFVEIINSTHASVYEAGPCWV (SEQ ID NO:324). An

additional preferred polypeptide fragment of the invention comprises the following amino acid sequence: AGIRHERNRGRLLCMLALTFMFMVLEVVSRR
VTSSLAMLSDSFHM LSDVLALVVALVAERFARRTHATQKNTFGWIRAEVMGALVNAIFLTGLCFAILLE
AIERFIEPHEMQQPLVVLGVGVAGLLVNVVLGLCLFHHHSGFSQDSGHXHSHGHHGHLPGKPRVKST
RPGSSDINVAPGEQGPDQEETNTLVANTSNSNGLKLDPADPENPRSGDTVEVQVNGNLVREPDHMELEE
DRAGQLNMRGVFLHVLGDALGSVIVVNALVFYFSWKGCSEGDFCVNPCFPDPCKAFVEILIVLMHQFM
(SEQ ID NO:325). Polynucleotides encoding this sequence are also provided.

This gene is expressed primarily in colon, lung, liver, lymphoma, osteosarcoma, adrenal gland tumor and fibroblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders, as well as gastrointestinal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, gastrointestinal, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 156 as residues: Arg-50 to Thr-58, Ser-125 to Gly-132. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to ZnT-1 indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders associated with the regulation of zinc homeostasis. Although zinc is an important trace element in many biological systems, several lines of evidence suggest that this transporter may serve as a point of intervention particularly in the treatment of neurological diseases. The metabolism of zinc in the brain has been shown to be regulated by a number of transport proteins, including ZnT-1. Pharmacological doses of zinc cause neuronal death, and some estimates indicate that extracellular concentrations of zinc could reach neurotoxic levels under pathological conditions. In Alzheimer's disease, zinc has been shown to aggregate beta-amyloid, a form which is potentially neurotoxic. The zinc-dependent transcription factors NF-kappa B and Sp1 bind to the promoter region of the amyloid precursor protein (APP) gene. Zinc also inhibits enzymes which degrade APP to nonamyloidogenic peptides and which degrade the soluble form of beta-amyloid. The changes in zinc metabolism which occur during oxidative stress is important in neurological diseases where oxidative stress is implicated, such as Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis (ALS). Zinc is a structural component of superoxide dismutase 1, mutations of which give rise to one form of familial ALS. After HIV infection, zinc deficiency is found which is secondary to immune-induced cytokine synthesis. Zinc is involved in the replication of the HIV virus at a number of sites. Collectively, this transporter may prove useful in the treatment and diagnosis of several disorders related to zinc regulation. Alternatively, the tissue distribution within lymphomas indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in immune tissue indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or

other processes that may also suggest a usefulness in the treatment of cancer e.g. by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1882 of SEQ ID NO:37, b is an integer of 15 to 1896, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 28

The translation product of this gene was shown to have homology to the mouse interferon-stimulated gene 15 and human calnexin (See Genbank Accession Nos. gb|AAB02697.1| and gi|306481|gb|AAA21013.1|; all references available through these accessions are hereby incorporated by reference herein) which may implicate this gene as playing a role in regulation of proliferating and differentiating cells.

Preferred polypeptides comprise the following amino acid sequence:

MFTFASMTKEDSKLIALIWPSEWQMIQKLFVVDHVIKITRIEVDVNPSETQYISEPKLCPECREGLLC
 QQRDLREYTQATIYVHKVVDNKKVMDKSAPELNVSSSETEEDKEEAKPDGEKDPDFNQSXGGTKRQKI
 SHQNYIAYQKQVIRRSRHRKVRGEKALLVSANQTLKELKIQIMHAFSVAPFDQNLSIDGKILSDDCAT
 5 LGTLGVIPESVILLKADEPIADYAAMDVQMVCMPPEGFKGTGLLGH (SEQ ID NO:326);
 SAPELNVSSSETEEDKEEAKP (SEQ ID NO:327);
 FQDKNRPCLSNWPEDTDVLYIVSQFFVEEWRKFVRKPTRCSPVSSVGNSALLCPHGGL (SEQ ID
 NO:329); MFTFASMTKEDSKLIALIWPSEWQMIQKLFVVDHVIKITRIE (SEQ ID NO:330);
 VGDVNPSETQYISEPKLCPECREGLLCQQRDLREYTQATIY (SEQ ID NO:331); VHKVVDNK
 10 KVMKDSAPELNVSSSETEEDKEEAKPDGEKDPDF (SEQ ID NO:332); NQXGGTKRQKISHQN
 YIAYQKQVIRRSRHRKVRGEKALLV (SEQ ID NO:333); SANQTLKELKIQIMHAFSVAPFDQ
 NLSIDGKILSDDCATLGT (SEQ ID NO:334); LGVIPESVILLKADEPIADYAAMDVQMVCMP
 PEGFKGTGLLGH (SEQ ID NO:335); and/or KELKIQIMHAFSVAPFDQ (SEQ ID
 NO:328). Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed primarily in brain and hematological tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, cancers, developmental and regulatory diseases of the brain and
 20 immune system. Similarly, polypeptides and antibodies directed to these polypeptides
 are useful in providing immunological probes for differential identification of the
 tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
 particularly of the brain and immune system, expression of this gene at significantly
 higher or lower levels is routinely detected in certain tissues or cell types (e.g.,
 25 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
 synovial fluid and spinal fluid) or another tissue or cell sample taken from an
 individual having such a disorder, relative to the standard gene expression level, i.e.,
 the expression level in healthy tissue or bodily fluid from an individual not having the
 disorder.

30 Preferred polypeptides of the present invention comprise immunogenic
 epitopes shown in SEQ ID NO: 157 as residues: His-26 to Phe-31. Polynucleotides
 encoding said polypeptides are also provided.

The tissue distribution in brain indicates that polynucleotides and polypeptides
 corresponding to this gene are useful for the detection, treatment, and/or prevention of

neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of

5 Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS,

10 psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, expression in T-cells and bone marrow, and homology to the mouse interferon-stimulated gene 15 and human calnexin proteins indicate that the protein product of this gene might also be useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-

15 immunities, immunodeficiencies (e.g., AIDS), immuno-suppressive conditions (transplantation) and hematopoietic disorders. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of general microbial infection, inflammation, and cancer (e.g., by boosting immune responses). Furthermore, the protein may also be used to

20 determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

25 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

30 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1138 of SEQ ID NO:38, b is an

integer of 15 to 1152, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

5 Preferred polypeptides of the invention comprise the following amino acid sequence: RGERSEELLGREGLSGSQ (SEQ ID NO:336), and/or AEAAGEKEGVRSCWAER DCPAPRCWASWGAQPSWDGSQVLLWRSCCCCCWPPAFSTDGRTVTWRGTVQLQGETESAGPSLGPSSG GATWESFTITVILATYLMCRMWASTTTTTTPATXLTTXTTTTTTPTATIPATLAEAAVAGACGQQLPLPSH LFPQGQVDFMPCGRMHLWGERXEQ (SEQ ID NO:337). Polynucleotides encoding these
10 polypeptides are also provided.

This gene is expressed primarily in placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, developmental anomalies or fetal deficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing fetus, expression of this gene at significantly higher or lower levels is routinely detected in certain
20 tissues or cell types (e.g., reproductive, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid, amniotic fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 158 as residues: Gly-35 to Asp-40, Asn-51 to Trp-59. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in placenta indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of
30 developmental anomalies or fetal deficiencies, reproductive dysfunction, as well as ovarian and other endometrial cancers. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate

ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1003 of SEQ ID NO:39, b is an integer of 15 to 1017, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

The translation product of this gene shares sequence homology with ALS (Acid Labile Subunit of Insulin-Like Growth Factor) which is thought to be important in the regulation of IGF availability. As such, it is likely that the product of this gene
20 is involved in the regulation of various proliferation-dependent cellular processes that is attributable to cancer progression (See Genbank Accession No. gi|184808; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid
25 sequence: FHGLGRLHTVHL (SEQ ID NO:338), AAFGLALLEQLDLSDNAQLR (SEQ ID NO:339), HEVPDAPRPTPT (SEQ ID NO:341), and/or AFRGLHSLD (SEQ ID NO:340). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 22. Accordingly, polynucleotides related to this invention are useful as a marker in
30 linkage analysis for chromosome 22.

This gene is expressed primarily in cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases, growth deficiencies, osteoporosis, catabolic disorders and diabetes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system and other periferial tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, proliferating, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 159 as residues: Thr-41 to Gly-47, Pro-170 to Asp-176, Leu-257 to Trp-262, Gln-276 to Ser-283, Arg-323 to Leu-330, Pro-362 to Val-374. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution cerebellum and homology to ALS (Acid Labile Subunit of Insulin-Like Growth Factor) indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of a variety of metabolic disorders, growth deficiencies, osteoporosis, catabolic disorders (including AIDS) and diabetes. Nearly all of the insulin-like growth factor (IGF) in the circulation is bound in a heterotrimeric complex composed of IGF, IGF-binding protein-3, and the acid-labile subunit (ALS). The protein product of this gene therefore may afford the ability to potentiate the biological actions of IGF or similar growth factors and cytokines. Studies which demonstrate the beneficial effect of IGF-I in amyotrophic lateral-sclerosis, would suggest a role in this disease as well. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1763 of SEQ ID NO:40, b is an integer of 15 to 1777, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene was shown to have homology to diacylglycerol kinase which is known to be important in lipid metabolism (See Genebank Accession No.gi|1939; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: MVDADRNRASSSSYLCLLLFSLSLFLCHETVCDRATCLFFFLKFFFLFMCRCMSW GFKNFKAGLLMQSMPTSGILRERKRLHVVRIPQGTEKKLETVMQI (SEQ ID NO:342), and/or IPQGTEKKLETV (SEQ ID NO:343). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and neurodegenerative diseases of the brain and nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels

is routinely detected in certain tissues or cell types (e.g., neural, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or
5 bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 160 as residues: Gly-49 to Ser-54, Lys-61 to Arg-68. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain combined with the homology to a known
10 enzyme involved in lipid metabolism indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly,
15 the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive
20 compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In particular, this gene may have utility in the diagnosis, treatment, and/or prevention of disorders involving the PNS, CNS and/or other tissues which rely on lipid-containing structures such as myelin sheath
25 dependent nerves. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 989 of SEQ ID NO:41, b is an integer of 15 to 1003, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

10 This gene is expressed primarily in amygdala.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and neurodegenerative diseases of the brain and nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 161 as residues: Met-1 to Lys-6. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in amygdala indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection,

treatment, and/or prevention of aphasia, depression, schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, specific brain tumors, mania, dementia, paranoia, addictive behavior and sleep disorders. The amygdala processes sensory information and relays this to other areas of the brain including the endocrine and autonomic domains of the hypothalamus and the brain stem. As such, The translation product of this gene may show commercial utility in the diagnosis, treatment, and/or prevention of various endocrine, cardiovascular, and pulmonary disorders, particularly those disorders directly associated with CNS/autonomic control. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1187 of SEQ ID NO:42, b is an integer of 15 to 1201, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 33

The gene encoding the disclosed cDNA is believed to reside on chromosome 9. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 9.

Preferred polypeptides of the invention comprise the following amino acid sequence: NPRLPLPRGGSLRLLSSPANSNNAKAYPFSRFPSPIF (SEQ ID NO:344). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in B-cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, haemopoietic and immune diseases and/or disorders including cancer.

5 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemopoietic and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune,
10 hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15 The tissue distribution in B-cell lymphoma indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune
20 Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy,
25 immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis,
30 granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and

graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma.

Furthermore, the protein may also be used to determine biological activity, to raise

5 antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
10 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1162 of SEQ ID NO:43, b is an integer of 15 to 1176, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

20 **FEATURES OF PROTEIN ENCODED BY GENE NO: 34**

This gene is expressed primarily in breast cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, diseases and/or disorders of the reproductive organs and cancer, particularly of the mammary glands. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at
30 significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, breast, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell

sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
 5 epitopes shown in SEQ ID NO: 163 as residues: Asp-77 to Gly-127. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in tumors of breast origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of such tumors, in addition to other tumors. Representative uses are
 10 described in the "Hyperproliferative Disorders", "Infectious Disease", and "Binding Activity" sections below, in Example 11, and 27, and elsewhere herein. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
 15 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of
 20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 555 of SEQ ID NO:44, b is an
 25 integer of 15 to 569, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

Preferred polypeptides encoded by this gene comprise the following amino
 30 acid sequence: MVQEAPALVRLSLGSHRVKGPLPVLKLQPEGWSPSTLWSCASVWKDSC (SEQ ID NO:345), and/or ALASSLVAENQGFVAALMVQEAPALVRLSLGSHRVKGPLPVLKLQPEGWSPST

LWSCASVWKDSCMHPWRLSMCPACVLAALPALCSCLCSPDARPPHGWMSMPFTPHPLVSRAMPTCHPCS
(SEQ ID NO: 346). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in
5 linkage analysis for chromosome 11.

This gene is expressed primarily in placenta, dendritic cells, brain, and to a lesser extent in infant cells and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of developing cells and tissues, particularly growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
15 tissues or cells, particularly of the placenta and other developing organs and tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, neural, placental, brain, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
20 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 164 as residues: Pro-27 to Gly-34. Polynucleotides
25 encoding said polypeptides are also provided.

The tissue distribution in placental tissue indicates the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product
30 may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis.

Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. The expression within cellular sources marked by proliferating cells (e.g., infant cells and tissues) indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 972 of SEQ ID NO:45, b is an integer of 15 to 986, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

The translation product of this gene shares sequence homology with ion channel proteins which are thought to be important in many physiological processes including neural and muscular function (See, for example, Genebank Accession No. gi|1065507, and gb|AAC68885.1; all references available through these accession numbers are hereby incorporated herein; for example, FEBS Lett. 445, 231-236 (1999)). Specifically, this protein is homologous to the putative four repeat ion channel of *Rattus norvegicus*. Based upon the sequence similarity, The translation product of this gene is expected to share at least some biological activities with ion channel proteins. Such activities are known in the art, some of which are described elsewhere herein.

Preferred polypeptides comprise the following amino acid sequence:
FYFITLIFFLAWLVKNVFIIVIIETFAEIRVQF (SEQ ID NO:347), SIFTVYEASQEGWV (SEQ ID NO:348), and/or HEGTSIFTVYEASQEGWVFL (SEQ ID NO:349). Also preferred are polynucleotides encoding these polypeptides.

This gene is expressed primarily in spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the central and peripheral nervous system, particularly neural degenerative conditions, and is useful in restoring cognitive function. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neural system, expression of this gene at significantly higher or lower levels is

5 routinely detected in certain tissues or cell types (e.g., neural, brain, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 165 as residues: Phe-8 to Ser-13, Ala-84 to Ser-90. Polynucleotides encoding said polypeptides are also provided.

10 The tissue distribution in spinal cord tissue, combined with the homology to ion channel proteins, indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but
15 are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive
20 disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition,
25 homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
30 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
5 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1526 of SEQ ID NO:46, b is an integer of 15 to 1540, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

10 **FEATURES OF PROTEIN ENCODED BY GENE NO: 37**

When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the early growth response gene 1 (EGR) pathway. Thus, it is likely that this gene activates fibroblast cells, and to a lesser extent, other cells and tissue cell-types, through the EGR signal transduction pathway. The early
15 growth response gene is a separate signal transduction pathway from the Jaks-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in uterus, colon cancer, synovium, fetal lung, and to a lesser extent in fetal and adult heart.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders of developing cells and tissues, particularly infertility and cancer. Similarly, polypeptides and antibodies directed to these
25 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing and reproductive systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, developing, gastrointestinal, synovium, skeletal,
30 heart, lung, cardiovascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative

to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 166 as residues: Lys-32 to His-38. Polynucleotides
5 encoding said polypeptides are also provided.

The tissue distribution in developing and reproductive tissues, combined with the detected EGR1 biological activity, indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and
10 other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers,
15 or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type
20 specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to certain types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.
25 The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents
30 that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 778 of SEQ ID NO:47, b is an integer of 15 to 792, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

Preferred polypeptides of the invention comprise the following amino acid sequence: CKTSFGLA (SEQ ID NO:350). Polynucleotides encoding these polypeptides are also provided. In an alternative embodiment, polypeptides of the invention comprise the following amino acid sequence: MITLSSAFSAKQKTHAHKNTHACMCATDMANPKLVLFHFEVIVALLSLLQTLISLLLGQRTWLALHLYVLSTENXALHTVGTQKHLLPHDWCFCGKHCVSCRHHIFHRFCSIFSSTLKRSQGFEG (SEQ ID NO:351). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in fetal bone, B and T cell lymphoma, and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic, skeletal, and immune diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, skeletal, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
5 epitopes shown in SEQ ID NO: 167 as residues: Ser-33 to His-42. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in T-cells and dendritic cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia,
10 thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or
15 chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Moreover, the protein
20 may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
25 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1483 of SEQ ID NO:48, b is an integer of 15 to 1497, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

This gene is expressed primarily in prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive diseases and/or disorders, particularly prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, prostate, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 168 as residues: Pro-21 to Pro-26, Arg-31 to Asn-37. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in prostate tissue indicates that the protein products of this gene are useful for the diagnosis and intervention of prostate cancers, in addition to other tumors within the urogenital and reproductive system. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions,

in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1326 of SEQ ID NO:49, b is an integer of 15 to 1340, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 40

The translation product of this gene shares sequence homology with the human proliferating-cell nucleolar antigen as well as to a protein from *Schizosaccharomyces pombe* of unknown function (See Genebank Accession Nos. 189422 and gnl|PID|e349594, as well as Medline Article 90315275; all references available through these accessions are hereby incorporated herein by reference). This protein is the most cancer specific of the proliferation- associated nucleolar proteins identified thus far. In addition, it is of special interest because of its expression pattern in the early G1 phase, and, in studies prior to 1989, it has not been detected in benign tumors and most normal resting tissues.

25 In another embodiment, polypeptides of the invention comprise the following amino acid sequence:

SATEHGAVCCSCRRVGRGEPPGSIKGLVYSSNFQNVKQLYALVCETQRYSAVLDAVIASAGLL
RAEKKLRPHLAKVLVYELLGKGFRGGGRWKALLGRHQARLKAELARLKVHRGVSERNEDLLEVGSRPG
P
30 ASQLPRFVRVNTLKTCSDDVDYFKRQGF SYQGRASSLDDLRLKGKHFLDPLMPELLVFPAQTDLHE
H
PLYRAGHLILQDRASCLPAMLLDPPGSHVIDACAAPGNKTSHLAALLKNQKIFAFDLDAKRLASMAT
L

LAXAGVSCCELAEDFLAVSPXDPRYXEVHYXLLDPSCSGSGMPSRQLEXPGAGTPSPVRLHALAGFQQ
 RALCHALTFFPSLQRLVYSTCSLCQEENEDVVRDALQQNPGAFRLAPALPAWPHRGLSTFPGA EHCLRAS
 PE TTLSSGFFVAVIERVEXPSSASQAKASAPERTPSPAPKRKKRQQRAAAGACTPPCT (SEQ ID
 NO:356), CAAPGNKTSHLAA (SEQ ID NO:352), EHPLYRAGHLILQDRASCLPAMLL (SEQ
 5 ID NO:353), LLDPSCSGSGMPSRQ (SEQ ID NO:354), YSTCSLCQEENEDVVRDALQQNP
 (SEQ ID NO:355), and/or YEPHSTHSRERAMTSHARVSLGPSRDPLERPHLAKVLVYELLGK
 GFRGGGGRWKALLGRHQARLKAELARLKVHRGVSRNEDLLEVGSRP GPASQLPRFVRVNTLKTCSDDVV
 DYFKRQGSYQGRASSLDDL RALKGKHFLLDPLMPELLVFPAQTDLHEHPLYRAGHLILQDRASCLPAM
 LLDPPPGSHVIDACAAPGNKTSHLAALLKNQKIFAFDLDAKRLASMATLLAXAGVSCCELAEDFLAV
 10 SPXDPRYXEVHYXLLDPSCSGSGMPSRQLEEPGAGTPSPVRLHALAGFQQRALCHALTFFPSLQRLVYST
 CSLCQEENEDVVRDALQQNPGAFRLAPALPAWPHRGLSTFPGA EHCLRAS PETTLSSGFFVAVIERVEV
 PSSASQAKASAPERTPSPAPKRKKRQQXAAAGACTPPCT (SEQ ID NO:357).

Polynucleotides encoding these polypeptides are also provided. This gene maps to
 chromosome 7, and therefore, is used as a marker in linkage analysis for chromosome
 15 7.

This gene is expressed primarily in T cells and rejected kidney and to a lesser
 extent in keratinocytes and various other normal and transformed, predominately
 haemopoietic cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as
 20 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, immune diseases and/or disorders, particularly host-vs-graft disease,
 and transplant rejection. Similarly, polypeptides and antibodies directed to these
 polypeptides are useful in providing immunological probes for differential
 25 identification of the tissue(s) or cell type(s). For a number of disorders of the above
 tissues or cells, particularly of the immune system, expression of this gene at
 significantly higher or lower levels is routinely detected in certain tissues or cell types
 (e.g., rejected transplant tissue, immune, haemopoietic, and cancerous and wounded
 tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal
 30 fluid) or another tissue or cell sample taken from an individual having such a
 disorder, relative to the standard gene expression level, i.e., the expression level in
 healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells and rejected kidney, indicates
 polynucleotides and polypeptides corresponding to this gene are useful for the

diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1525 of SEQ ID NO:50, b is an integer of 15 to 1539, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

This gene is expressed primarily in placenta, uterus, 12 week old, early stage, embryo and to a lesser extent in epithelium.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and reproductive diseases and/or disorders, in addition to disorders of the integumentary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developmental and epithelial tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developmental, reproductive, uterine, placental, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in placental, uterine, and embryonic cells and tissues indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections

below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is
5 believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and
10 polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to
15 aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. The protein is useful for the detection, treatment, and/or prevention of various types of cancer, particularly of the integumentary system. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
20 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1409 of SEQ ID NO:51, b is an

integer of 15 to 1423, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

5 The translation product of this gene was shown to have homology to the human, bovine, mouse, and rat G protein gamma-3 subunit (See Genebank Accession Nos. W09413, pir|A36204|RGBOG3, gi|2582400 (AF022088), and gi|1353498) which are known to play a role in the regulation of signal transduction pathways. Moreover, the protein shares structural homology to a yeast mitochondrion membrane protein
10 Q0225 (See Genbank Accession No. pir|S72689|S72689).

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

15 NREQKAKSQLLRSQLYSTLDLPYFFQCVGTRCTAVCVCVCVCVCVCX
YLPIHWQVNLHLVYLAMLCFLPIPLLSILSPQTQASRLLEDTVRRKHFLTYPFG
ISSIITQALL (SEQ ID NO:360). Polynucleotides encoding these polypeptides are also provided.

In yet another embodiment, polypeptides of the invention comprise the
20 following amino acid sequence: MGTHSVSGRFSKTSPPYCPFSSSLPGPISSIGFNKSLHECL
FISEKELLPLFPFPDLKSFISYLTSMKPGPLIVSLKIWVSYPITRPRYLPPMLKSLNISFLYIQYIW
AYIHLYSFYIYIISVSFFLDKPFYIYISFPPKPHFLFASLSKTQEFHFHVPQHFFFLIFSPQVSSPIS
CFARLLKSPLFTPVPTEISPFYNCAYYADIPSPQLVWGPISHQTWLLKLGLLPKRGFQVRGDRL
(SEQ ID NO:358), and/or CFARLLKSPLFTPVPTEISPFYNCAYYSA (SEQ ID
25 NO:359). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain, fetal tissue, frontal cortex, corpus collosum, and to a lesser extent in amygdala tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
30 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural and CNS diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological

probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and peripheral nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, and cancerous
5 and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
10 epitopes shown in SEQ ID NO: 171 as residues: Thr-26 to Leu-33. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in various neural cells and tissues, combined with the similarity to G Protein Gamma-3 subunit indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of
15 neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome,
20 meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep
25 patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as
30 tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,

antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1350 of SEQ ID NO:52, b is an integer of 15 to 1364, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

The translation product of this gene shares homology with the human alpha-3 type IX collagen protein (See Genebank Accession No.gi|1196421). This protein likely represents a Type IIIb membrane protein. Although the preferred open reading frame of the present invention contains a signal peptide (as delineated in Table 1 and described elsewhere herein), the protein appears to have several transmembrane domains. The transmembrane domains are located at about amino acid position 111 - 162, 137 - 162, 163 - 186, and 64 - 85 of the sequence referenced in Table 1 for this gene. Preferred are polypeptides comprising the following amino acid sequence: PGPEAQPWPGPDLPA VGSRGPGRLLAAVSAPRLGLGLAGADPVGPEACHLP (SEQ ID NO: 361), GRLRGPDEVGAPFHPGPATPGLADPLRPAEPXHWLPSLWGPT (SEQ ID NO: 362), PGPEAQPWPGPDLPAVGSR (SEQ ID NO: 363), and/or ATPGLADPLRPAEPXHWLP (SEQ ID NO: 364). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

QWPEKDPVMAASSISSPWGKHVFKAILMVLVALILLHSALAQSRDFAPP
GQQKREAPVDVLTQIGRSVRGTLDAWIGPETMHLVSESSSQVLWAISSAISVAFFALSGIAAQLLNALG
LAGDYLAQGLKLSPGQVQTFLLWGAGALVVYWLSSLGLVLALLGRILWGLKLVIFLAGFVALMRSVP

DPSTRALLLLALLLILYALL SRXTGSRASGAQLEAKVRGLERQVEELRWRQRQXAKGARSVEEE (SEQ ID NO:365) . Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

This gene is expressed primarily in melanocytes, and to a lesser extent in synovial sarcoma and larynx sarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, melanoma and other disorders of the integumentary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the synovial and epithelial tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 172 as residues: Gln-15 to Phe-20, Pro-22 to Ala-30, Val-160 to Thr-165. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in melanocytes and sarcoma tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study treatment and diagnosis of various cancers and their metastases, particularly of the integumentary system. Additionally, the homology to a conserved collagen protein would suggest that this protein may also be important in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal

chondrodysplasia type Schmid. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", "Infectious Disease", and "Regeneration" sections below, in Example 11, 19, and 20, and elsewhere herein. Briefly, the protein is useful in detecting, treating, and/or preventing congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders (i.e., arthritis, trauma, tendonitis, chondromalacia and inflammation, etc.), autoimmune disorders (i.e., rheumatoid arthritis, lupus, scleroderma, dermatomyositis, etc.), dwarfism, spinal deformation, joint abnormalities, and chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2274 of SEQ ID NO:53, b is an integer of 15 to 2288, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

The translation product of this gene shares sequence homology with tumor progression inhibitor which is thought to be important in inhibition of tumor growth as well as its metastasis (See Genebank Accession No. W26667). Preferred are polypeptides comprising the following amino acid sequence:

EXPRXIXGXNAPQVPVRNSR
VDPRVRPRVRSLVFVLCDEVQRQWYVNGVNYFTDLWNVMDTLGLFYFIAGIVFRLHSSNKSSLYSGRVI
FCLDYIIFTLRILHIIFTVSRNLGPKII (SEQ ID NO:366), NILLVNLLVAMF (SEQ ID
NO:367), and/or QVWKFORFYFL (SEQ ID NO:368). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

EXPRXIXGXNAPQVPVRNSRVDPRVRPRVRSLVFVLCDEVQRQWYVNGVNY
FTDLWNVMDTLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRILHIIFTVSRNLGPKIIMLQR
MLIDVXXFLFLFAVWMVAFGVAXQGILRQNEQRWRWIFRSVIYEPXLAMFGQVPSXVDGTTYDFAHCTF
TGNEKPLCVXLDEHNLPRFPWITIPLVCIYMLSTNILLVNLLVAMFGYTVGTQENNDQVWKFORFYF
LVQEYCSRLNIPFPFIVFAYFY MVVKCKFCCKCKEXNXESSVCCSKMXTMRLWHGRVS (SEQ ID
NO:369). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in adult liver, prostate, gall bladder, and to a lesser extent, in Hodgkin's lymphoma II.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, liver cancer and other hepatic diseases and/or disorders. Similarly,

polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, expression of this gene at significantly higher or lower levels is routinely detected in certain
5 tissues or cell types (e.g., hepatic, reproductive, metabolic, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, bile, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
10 having the disorder.

The tissue distribution in liver and gall bladder cells and tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers. Representative uses are described in the "Hyperproliferative Disorders", "Infectious Disease", and "Binding
15 Activity" sections below, in Example 11, and 27, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, and/or prevention of hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of
20 various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
25 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1498 of SEQ ID NO:54, b is an integer of 15 to 1512, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 45

The polypeptide of the present invention is thought to have an intramitochondrial signal indicating that the protein could play a role in metabolic processes, including apoptosis. Based upon this fact, it is expected that the protein product of this gene will share at least some biological activities with other
10 mitochondrial proteins having a similar signal. Such activities are known in the art, some of which are described elsewhere.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the
15 following amino acid sequence:
MEFQNMYYIQLFGFSFFIVIVRMLLLGLCVSARQPVMPRATLWGHLSPA
WVLVPWTPRACGQAAPGRGHVASDHKSGLPWPKHCSCLHPRASQPCLFSLNSNRTVFTAIQRVALGWTF
WVQANLVPRCT (SEQ ID NO:370). Polynucleotides encoding these polypeptides are also provided.

20 The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed primarily in human prostate cancer, and to a lesser extent in soares melanocyte and human colon.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancer, melanoma, and other diseases and/or disorders of the integumentary system. Similarly, polypeptides and antibodies directed to these
30 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene

at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., prostate, reproductive, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, seminal fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
5 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 174 as residues: Ser-36 to Gly-41, Pro-43 to Ser-49. Polynucleotides encoding said polypeptides are also provided.

10 The tissue distribution in tumors of prostate, colon, and integument origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Representative uses are described elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or
15 prevention of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a
20 nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
30 formula of a-b, where a is any integer between 1 to 1343 of SEQ ID NO:55, b is an integer of 15 to 1357, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

LLLCVTGVSYGLMHPIPSSFMKAVSSFLTAEASVGNPEGAFMKVLQAR
 KNXTSTELIVEPEEPSDSSGINLSGFGSEQLDTNDESXISTLSYILPYFSAVNLDVXSXLLPFIKLPT
 XGNSLAKIQTVGQNXQXVXRVLGMGRSIQKRHFKEVGRQSIIRREQGAQASVENAAEEKRLGSPAPREXE
 QPHTQQGPEKLAGNAXYTKPSFTQEHKAAVSVLXPFSKGAPSTSSPAKALPQVRDRWKDXTHXISILES
 AKARVTNMKASKPISHSRKKYRFHKTRSRMTHRTPKVKKSPKFRKKSYLSRLMLANRPPFSAAXSLINS
 PSQGAFFSSLDLSPQENPFLXVSAPSEHFIETTNIKDTTARNALEENVFMENTNMPEVTISENTNYNHP
 PEADSGXTAFNLGPTVKQTET (SEQ ID NO:371). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in duodenum and cheek carcinoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointesinal disorders and carcinomas, in addition to disorders of the epithelium and mucosa. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., gastrointestinal, epithelial, mucosa, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in duodenal tissues and epithelia indicates that the protein product of this gene is useful for the diagnosis and intervention of tumors and other disorders within these tissues, in addition to other tumors. The expression within embryonic tissue and other cellular sources marked by proliferating cells indicates

this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1975 of SEQ ID NO:56, b is an integer of 15 to 1989, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of this gene shares sequence homology with mouse magnesium dependent protein phosphatase (See Genebank Accession Nos. gnl|PID|d1004752 and emb|CAA06555.1| (AJ005458); all references available through these accessions are hereby incorporated herein by reference; for example, J. Neurosci. Res. 51 (3), 328-338 (1998)) which is thought to be important in normal

protein metabolism and possibly gene regulation. Based on the sequence similarity,

The translation product of this gene is expected to share at least some biological activities with phosphatase proteins. Such activities are known in the art, some of which are described elsewhere herein.

5 Preferred polypeptides comprise the following amino acid sequence:

CFSNAPKVSDEAVKKDSELDKHLESRVEEIMEKSGEEGMPDLAHVMRILSAENIPNLPPGGGLAGXRV
IEAVYSRLNPHRESGGAGDLEDPW (SEQ ID NO: 372), CFSNAPKVSDEAVKKDSELDKHLES
RVEEIMEKSGEEGMPDLAHVMRILSAENIPN (SEQ ID NO: 373), RNVIEAVYSRLNPHRESGG
GAGDLED (SEQ ID NO: 374), DSELDKHLESRVEEIM (SEQ ID NO: 375), KSGEEGMP
10 DLAHVMRILSAENIPN (SEQ ID NO: 376), and/or CFSNAPKVS (SEQ ID NO: 377).
Polynucleotides encoding these polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following
amino acid sequence: MSRKSLAFPIICSYLCLTVATCSIACTTVFFANLRHTRYICIELSALET
SGVISPQINNVPVHGKYS (SEQ ID NO: 378). Polynucleotides encoding these
15 polypeptides are also provided.

This gene is expressed primarily in prostate and to a lesser extent in melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, proliferative conditions and cancers, in addition to reproductive, visual, and integumentary diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of
25 the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, visual, retinal, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, aqueous humor, vitreous humor, synovial fluid and spinal fluid) or another tissue or cell sample taken
30 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 176 as residues: Asp-6 to His-13, Asp-114 to Gly-131, Thr-166 to Gln-181, Val-210 to Thr-216, Pro-222 to Tyr-227. Polynucleotides encoding said polypeptides are also provided.

5 The tissue distribution in prostate tissue, combined with the homology to mouse magnesium dependent protein phosphatase indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study and treatment of various cancers and reproductive disorders. This protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment,
10 and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in
15 inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). This protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative
20 conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. The activity of this protein has been determined to be dependent upon the presence of magnesium ions. This protein is useful in the treatment,
25 detection, and/or prevention of various visual disorders, particularly degenerative conditions, and retinitis pigmentosa. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the
30 protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2529 of SEQ ID NO:57, b is an integer of 15 to 2543, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

The translation product of this gene shares sequence homology with ribosomal protein L32 and L14, a mitochondrial protein from rat tissues thought to be important in translation (See Genebank Accession No.gi|868267). Preferred are polypeptides comprising the following amino acid sequence: IQKMTRVRVVDNSALG (SEQ ID NO: 379), PRCIHVYKKNVGK (SEQ ID NO: 380), GDQILLAIKGQKKKA (SEQ ID NO: 381), and/or NPVGTRIKTPIPTSL (SEQ ID NO: 382). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

VLIPSFSSSFLCSRGGPLPXDLSDPMAFFTGLWGPFTCVSRVLSHHCF
STTGSLSAIQKMTRVRVVDNSALGNSPYHRAPRCIHVYKKNVGKVGQILLAIKGQKKKALIVGHCMF
GPRMTPRFDSNNVLIEDNGNPVGTRIKTPIPTSLRKREGEYSKVLAIQNFV (SEQ ID NO: 383). Polynucleotides encoding these polypeptides are also provided. This gene maps to chromosome 6, and therefore, is used as a marker in linkage analysis for chromosome 6.

This gene is expressed primarily in uterus, fetal liver/spleen, human endometrial stromal cells-treated with estradiol and amniotic cells - Primary Culture, and to a lesser extent in, human fetal kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometriosis and reproductive disorders, particularly of the female reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., uterine, endometrium, reproductive, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 177 as residues: Pro-92 to Ser-102, Leu-127 to Tyr-134. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in endometrium and uterine tissues, combined with the homology to a ribosomal protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of tumors within said tissue, in addition to other tumors where expression has been indicated. This protein may play a role in cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this

gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Antagonists, including antibodies directed against this invention, is useful in inhibiting cellular proliferation and thus is useful in inhibiting cancers, in addition to other proliferative diseases and/or disorders. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues.

10 The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show

15 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 763 of SEQ ID NO:58, b is an integer of 15 to 777, where both a and b correspond to the positions of nucleotide

25 residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

This gene is expressed primarily in liver, hepatoma and to a lesser extent in epithelial-TNF α and INF induced.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, liver diseases and/or disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, liver, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 178 as residues: Glu-28 to Gly-45, Ser-63 to Gly-69, Gln-96 to Trp-104, Gly-112 to Pro-117, Arg-121 to Pro-128. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in liver and hepatoma tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Representative uses are described in the "Hyperproliferative Disorders", "Infectious Disease", and "Binding Activity" sections below, in Example 11, and 27, and elsewhere herein. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
 5 formula of a-b, where a is any integer between 1 to 865 of SEQ ID NO:59, b is an integer of 15 to 879, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

10 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ARVVQPAARAGMWAGGRSSCQAEVLRATRGGGAARGNAAPGRALEMVPGAAG

15 WCCLVLWLPACVAAHGFRIDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQIHLVPAEPPEACGE
 LSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDNALMTASTWR (SEQ ID NO: 384).

Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome
 2. Accordingly, polynucleotides related to this invention are useful as a marker in
 20 linkage analysis for chromosome 2.

This gene is expressed primarily in breast lymph node, ovary, osteoclast cells, and to a lesser extent in human jurkat membrane-bound polysomes and human placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as
 25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer and immune diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
 30 a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, endocrine, skeletal, bone, placental,

and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in human breast and placental tissue indicates that the protein product of this gene is useful for diagnosis and intervention of tumors within these tissues, in addition to other tumors and tissues where expression has been indicated. Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1147 of SEQ ID NO:60, b is an integer of 15 to 1161, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

following amino acid sequence:

IATAALFFFFYCQVAGFIGKGQSLRSWVPQRLGLLEPQLQPMQSRLLLP
 FLFFLLEGCAPSSLGPGAAPGSGHSLGPPGSPGAPGPQPAVGPSPPCQPGPSPSSPAAAAASSQSSVAS
 WPCTLRCAAPSPDASALRPAASPAATPAWSPGSGTIRVLRPPAPAAAPATAITNRGPPRRRRRNARTA

5 (SEQ ID NO: 385). Polynucleotides encoding these polypeptides are also provided.

In yet another embodiment, polypeptides of the invention comprise the following amino acid sequence: ERPPPRRTGTPVARPRGPPDPAVAAGTALRAKQFARYGAASG
 VVPGSLWPSPEQLRELEAEEREWYPSLATMQESLRVKQLAEEQKRREREQHIAECMAKMPQMIVNWQQQ
 QRENWEKAQADKERRARLQAEAQELLGYQVDPRSARFQELLQDLEKKERNPQGGKTETEEGGATAALAA
 10 AVAQDPAASGAPSS (SEQ ID NO: 386). Polynucleotides encoding these polypeptides are also provided. The polypeptide sequence of the latter embodiment was found to have homology to the human HPK/GCK-like kinase HGK (See Genbank Accession No. gb|AAD16137.1| (AF096300); all references available through this accession are hereby incorporated herein by reference; for example, J. Biol. Chem. 274 (4), 2118-
 15 2125 (1999)) which is thought to play a role in modulating gene expression, particularly for genes involved in the c-jun pathway. Based on the sequence similarity, The translation product of this gene is expected to share at least some biological activities with signalling and kinase proteins. Such activities are known in the art, some of which are described elsewhere herein.

20 The gene encoding the disclosed cDNA is believed to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in HL-60, PMA 4H and to a lesser extent in Soares breast 2NbHBst, Human Pituitary, subt IX, and Human Fetal Kidney.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, developmental, and proliferative diseases and/or disorders, particularly promyelocytic leukemia. Similarly, polypeptides and
 30 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain

tissues or cell types (e.g., immune, hematopoietic, reproductive, developmental, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 180 as residues: Ser-54 to Ser-63, Asn-132 to Thr-145. Polynucleotides encoding said polypeptides are also provided.

10 The tissue distribution in HL-60 cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or
15 activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the
20 natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to
25 transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other
30 blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of

various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 673 of SEQ ID NO:61, b is an integer of 15 to 687, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

The translation product of this gene shares sequence homology with the human hypothetical L1 protein (third intron of gene TS) (See Genebank Accession No. pir|JU0033|JU0033), which is thought to be important for the regulation of RNA-dependent DNA polymerases.

Preferred polypeptides comprise the following amino acid sequence:

YQSLAETQKKENFRPISLKNTDAKILNKILANQIQQHIKKLIHNDRVGFIPMQGWFNICKSINIVHH
INRTKDKNHMIISIDAEKAFDKIRQSFMLKTLNKLGIHGYLGR (SEQ ID NO: 387), KKENFR
PISLKNTDAKILNKILANQIQQHIKKLIHNDRVGFIPMQGWFNICKSINIVHHINRTKDKNHMIISID
AEKAFDKIRQSFMLKTLNKLGIHGY (SEQ ID NO: 388), DAKILNKILAN (SEQ ID NO:
389), IQQHIKKLIH (SEQ ID NO: 390), KDKNHMIISIDAEKAFDKI (SEQ ID NO:
391), MLKTLNKLGI (SEQ ID NO: 392), and/or KKENFRPISL (SEQ ID NO:
393). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence: WTMFIDLHMLNQPCISGMKPTRSL

WISFLMCCWIWFANILLRIFASVFFRDIGLKFSFFCCVSARLWYQDDAGLINEL
GRIPSFY (SEQ ID NO: 394). Polynucleotides encoding these polypeptides are also
provided. The presence of the amino acid sequences upstream of the predicted signal
sequence of the latter embodiment may alter the characteristics of the protein of the
5 present invention such that either the full protein, or fragments thereof, are bound to
the membrane in a form analagous to a Type II membrane protein. This form of the
protein is thought to have a cytoplasmic tail covering about the first 21 amino acids.
Based on the structural similarity, the translation product of this latter embodiment is
expected to share at least some biological activities with type II membrane proteins.

10 Such activities are known in the art, some of which are described elsewhere herein.

This gene is expressed primarily in ulcerative colitis.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, gastrointestinal diseases and/or disorders, particularly ulcerative colitis.
Similarly, polypeptides and antibodies directed to these polypeptides are useful in
providing immunological probes for differential identification of the tissue(s) or cell
type(s). For a number of disorders of the above tissues or cells, particularly of the
digestive system, expression of this gene at significantly higher or lower levels is
20 routinely detected in certain tissues or cell types (e.g., gastrointestinal, and cancerous
and wounded tissues) or bodily fluids (e.g., lymph, chyme, bile, serum, plasma, urine,
synovial fluid and spinal fluid) or another tissue or cell sample taken from an
individual having such a disorder, relative to the standard gene expression level, i.e.,
the expression level in healthy tissue or bodily fluid from an individual not having the
25 disorder.

The tissue distribution in ulcerative colon tissue combined with its homology
to an RNA-dependent DNA polymerase regulatory protein may suggest that
polynucleotides and polypeptides corresponding to this gene are useful for diagnosis
and intervention of tumors and other proliferative conditions within the indicated
30 tissues, and to a lesser extent in other tissues and cell types. Moreover, the expression
within cellular sources marked by proliferating cells indicates this protein may play a
role in the regulation of cellular division, and may show utility in the diagnosis,

treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 504 of SEQ ID NO:62, b is an integer of 15 to 518, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 53

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

30 ERPEEGTEPSPSPVAEQASVSMTPVFRAWGLWVYVLPTGFPGPCCMMLLEL
FPKESVPQAYQGILLYLHFGF (SEQ ID NO: 395). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in ovary, testis, Hodkin's lymphoma, resting T-Cell; re-excision and to a lesser extent in soares multiple sclerosis, human corpus colosum, and fetal kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, immune, and hematopoietic diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
10 type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, ovarian, testicular, breast, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, seminal fluid, breast milk, plasma, urine, synovial
15 fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in testicular tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of
20 conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment
25 and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few
30 possible target indications. Moreover, the protein product of this gene has also been shown to be expressed in ovary and breast tissue which, in combination with the detected expression in testis, indicates that this protein represents a secreted factor

that plays an important role in proper reproduction (e.g., hormone, signalling factor, etc.). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional
5 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of
10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 897 of SEQ ID NO:63, b is an
15 integer of 15 to 911, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

When tested against U937 cell lines, supernatants removed from cells
20 containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other cells and tissue cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway
25 involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by
30 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence: RGE
VPHQPHPTRRTVVSGQAPWXP GPXALGQXVETAAGMGMP LVT VTAATFPTL

SCPPRAWPEVEAPEAPALP

VVPELPEVPMEMPLVLPPELELLSLEAVHRYQXGGTLMGWTRAEASANGS

(SEQ ID NO: 396). Polynucleotides encoding these polypeptides are also provided. In yet another embodiment,

- 5 Preferred polypeptides of the invention comprise the following amino acid sequence: MVLDPYRAVALELQANREPDFSSLSVSPSPRRMAARVFYLLLGECMHVCVCMWGRDTET RGPYRDSPDLSPRLLTALSATDSSRETRKAIWSPDPAGAQIPLRLESIYKAARKPATSSKPRRASL KKKKK (SEQ ID NO: 397). Polynucleotides encoding these polypeptides are also provided. Polypeptides of the latter embodiment share homology to the human
- 10 hHR21spB (See Genbank Accession No.gi|4101480|gb|AAD01193.1| (AF006264); all references available through this accession are hereby incorporated by reference herein) which is thought to play a role in DNA repair. Based on the sequence similarity, The translation product of this gene is expected to share at least some biological activities with DNA repair proteins. Such activities are known in the art,
- 15 some of which are described elsewhere herein.

The gene encoding the disclosed cDNA is believed to reside on chromosome 22. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 22.

- This gene is expressed primarily in resting T-Cells, testis, uterine cancer, bone
- 20 marrow, and to a lesser extent in cerebellum.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, reproductive, and neural diseases and/or disorders. Similarly,
- 25 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, neural, reproductive, and
- 30 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, seminal fluid, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in bone marrow and resting T-cells, combined with the detected GAS biological activity, indicates polynucleotides and polypeptides
5 corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or
10 activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an
15 agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host
20 diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene
25 product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Furthermore, the protein
30 may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,

antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 949 of SEQ ID NO:64, b is an integer of 15 to 963, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

The translation product of this gene was shown to have homology to the human platelet membrane glycoprotein V, which is a part of the Ib-V-IX system of surface glycoproteins (GPs Ib alpha, Ib beta, V, IX) that constitute the receptor for von Willebrand factor (vWf) and mediate the adhesion of platelets to injured vascular surfaces in the arterial circulation, a critical initiating event in hemostasis (See Genebank Accession No.gi|388760). Moreover, the protein product of this gene was also shown to have homology to human toll and toll-like receptors (See Genbank Accession Nos. W86352, and gb|AF051151|AF051151; all references available through this accession are hereby incorporated herein by reference; for example, Blood 91 (11), 4020-4027 (1998)). Based on the sequence similarity, The translation product of this gene is expected to share at least some biological activities with toll-receptor proteins. Such activities are known in the art, some of which are described elsewhere herein. Preferred are polypeptides comprising the following amino acid sequence: AFRNLPNLRIL (SEQ ID NO: 398), and/or AFQGLFHLFELRL (SEQ ID No: 399). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

NKXILEVPSARTTRIMGDHLDLLLGVVLMAGPVFGIPSCSFDGRIAFYR

FCNLTQVPQVLNLTTERLLLSFNYYIRTVTASSFPFLEQLQLLELGSQYTPLTIDKEAFRNLPNLRILDLG

5 SSKIYFLHPDAFQGLFHLFELRLYFCGLSDAVLKDG YFRNLKALTRLDLSKNQIRSLYLHPSFGKLNLSL

KSIDFSSNQIFLVCEHELE (SEQ ID NO: 400). Polynucleotides encoding these

polypeptides are also provided.

This gene is expressed primarily in pancreatic tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pancreatic cancer; impaired pancreatic function; altered carbohydrate metabolism; and immune and hematopoietic diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
15 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pancreas or endocrine system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., pancreatic, gastrointestinal, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g.,
20 lymph, serum, plasma, urine, bile, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in pancreatic tumors indicates that polynucleotides and
25 polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the pancreas. Expression of this gene product in pancreas tumors indicates a potential involvement in pancreatic cancer, and indicates that the gene product may play more general roles in cellular proliferation and/or apoptosis as well. Alternately, expression in the pancreas may suggest a general involvement in
30 pancreatic function, and implicate the utility of this gene product in a variety of pancreatic disorders. Alternately, as this protein is a secreted protein, it may simply be produced by the pancreas to have effects at other sites within the body or endocrine

system. In addition, the homology to a conserved receptor for von Willebrand factor indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. The product of this gene may also show utility in the treatment of vascular diseases such as atherosclerosis and stroke. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 987 of SEQ ID NO:65, b is an integer of 15 to 1001, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 56

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

AHAALQLSLRTCGPCSSPYPHAGLAALLTHMWALQLSLPTCGLAALLTHMRPCSSPYPHAGLAALLTHM
GPCRSPYPHGGLAAVLTHMRALQLSLPTWGLAALLTHMRPCSSPYPHAGLACCWLWSLSSHRSLSLQVQAT
5 HRLVVRTIKDRVMLKVLFPQTRRRGPFLLSSCRNDVMRNCVPRHAVLVTTTCVFVSFPTHCKVGITGPITQV
KQKPGNHSSPCPVIQLVAKAEFELMLPSVPKPVYLTLLVLSWCLCDVPCLSVSL (SEQ ID NO:

401). Polynucleotides encoding these polypeptides are also provided. It has been determined that the protein product of this gene has a conserved G-protein receptor motif beginning at amino acid position 89 and ending at amino acid position 105 of
10 the amino acid sequence referenced in Table 1 for this gene.

Preferred polypeptides of the invention comprise the following amino acid sequence: LACCWLWSLSSHRSLSLQV (SEQ ID NO: 402). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in tonsils and anergic T-cells.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders; immune dysfunction; impaired immune surveillance. Similarly, polypeptides and antibodies directed to these polypeptides are
20 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph,
25 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
30 epitopes shown in SEQ ID NO: 185 as residues: Pro-22 to Pro-28, Pro-41 to His-48, Pro-79 to His-86, Pro-126 to Phe-134, Ser-137 to Met-143, Gln-176 to Ser-186. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in T-cells and tonsils, combined with the identification of a G-protein receptor motif within the open reading frame, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the

5 "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or

10 other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease,

15 inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia,

20 rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of

25 various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
5 formula of a-b, where a is any integer between 1 to 1544 of SEQ ID NO:66, b is an integer of 15 to 1558, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

10 This gene is expressed primarily in healing groin wound (6.5 hours post incision), and to a lesser extent in testis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, wounded tissues; disorders involving tissue repair; male reproductive disorders; mucositis; tissue degeneration. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this
20 gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, testis, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in
25 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 186 as residues: Ser-59 to Gly-68. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in healing groin wound and testis indicates that
30 polynucleotides and polypeptides corresponding to this gene are useful for therapeutic use as an agent to facilitate wound healing and tissue regeneration. Expression of this product during wound healing indicates that it may play a beneficial role during the

process. Alternately, expression during wound healing may also suggest that it plays a negative role during the process, e.g. fibrosis and scarring, and that therapeutics designed to counter the effects of this protein is even more beneficial. In addition, expression of this protein within the groin and testis indicates that it may play a role in reproductive system function - particularly male reproductive function - and that this protein may even have potential uses as a male contraceptive. Alternately, The tissue distribution in testicular tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1308 of SEQ ID NO:67, b is an

integer of 15 to 1322, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

5 A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MGEASPPAPARRHLLVLLLLLSTLVIPSAAPLHDADAQESSLGLTGLOS LLQGFSRLFLKVTCFGA (SEQ ID NO: 403). Polynucleotides encoding these polypeptides are also provided.

10 This gene is expressed primarily in testis, and to a lesser extent in brain and fetal heart.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders; psychological disorders; learning
15 disabilities; altered heart function; altered male reproductive function. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and nervous system, cardiovascular system, or reproductive system, expression of this
20 gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, testis, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression
25 level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 187 as residues: Pro-82 to His-93. Polynucleotides encoding said polypeptides are also provided.

30 The tissue distribution in testicular tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment

of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene

5 expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Alternatively, The tissue distribution in brain indicates

10 that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of brain and nervous system disorders. Expression of this gene product in a variety of brain regions indicates a role in brain and nervous system function. This indicates that the protein product is useful in the treatment of neurodegenerative disorders; learning disabilities; psychoses; and behaviours,

15 including feeding; sleeping; perception; balance; etc. Therefore, this gene product is useful in the treatment of a variety of heart conditions, including myocardial infarction; congestive heart failure; arrhythmias; coronary occlusion; and a variety of other disorders of the heart. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or

20 receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell

25 proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for

30 treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction

etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 851 of SEQ ID NO:68, b is an integer of 15 to 865, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

The translation product of this gene shares sequence homology with alpha 1,3 galactosyltransferase which is thought to be important in the regulation of protein glycosylation and sugar transfer (See Genebank Accession No. bs|150271; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides comprise the following amino acid sequence:

MLVVSTVVIIVFEFINSTEGSFLWIIYHSKNPEVDDSSAQKGGWFLSWFNNGIHNYQQGEEDIDKEKGRE
ETKGRKMTQQSFGYGTGLIQT (SEQ ID NO: 404), and/or FPGRTHASGNVKGKVILS

(SEQ ID NO: 405). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ADQEKIRNVKGVILSMLVVSTVIIVFWEFINSTEGSFLWIYHSKNPEV

DDSSAQKGWFLSWFNNGIHNYQQGEEDIDKEKGREETKGRKMTQQSFGYGTGLIQT (SEQ ID NO:

5 406). Polynucleotides encoding these polypeptides are also provided. The presence of the upstream amino acids of the latter embodiment may significantly alter the secreted characteristics of the present invention. Namely, either the full-length protein, or fragments thereof, is some membrane bound in a mechanism analogous to type II membrane proteins. Based on the such characteristics, the translation product
10 of this latter embodiment is expected to share at least some biological activities with type II membrane proteins. Such activities are known in the art, some of which are described elsewhere herein. fragments.

The gene encoding the disclosed cDNA is believed to reside on chromosome 9. Accordingly, polynucleotides related to this invention are useful as a marker in
15 linkage analysis for chromosome 9.

This gene is expressed primarily in primary dendritic cells, neutrophils, and T cells and to a lesser extent in liver hepatoma and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune dysfunction, hematopoietic disorders; inflammation; neurodegenerative disorders; liver hepatoma; T cell lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
25 a number of disorders of the above tissues or cells, particularly of the immune system, liver, or CNS, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
30 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 188 as residues: His-27 to Gly-41, Gln-56 to Tyr-83. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in dendritic cells, combined with the homology to galactosyltransferases indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of disorders, particularly of the immune and nervous systems since normal function of such tissues depends upon proper glycoprotein recognition and galactosyltransferase function. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in dendritic cells indicates a role in the regulation of the immune system and responses to infectious agents. This may involve roles in antigen presentation, antigen processing, stimulation and activation of B and T cells, or stimulation/activation of dendritic cells themselves. This is evidenced by effects on cytokine production. Expression of this gene product in other hematopoietic cells such as T cells and neutrophils also indicates roles in the functions of those cells as well, and involvement in the proliferation, survival, and/or differentiation of hematopoietic cells in general. In addition, the expression also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses may include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. Expression of this gene product within infant brain also indicates a role in neuron survival, synapse formation, neurotransmission, perception, etc. The protein is useful in the treatment and/or prevention of degenerative myelinating diseases and/or disorders, particularly multiple sclerosis, in addition to other disorders which occur secondary to aberrant fatty-acid metabolism. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or

receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1136 of SEQ ID NO:69, b is an integer of 15 to 1150, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

This gene is expressed primarily in small intestine and leukocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders; inflammation; allergy; impaired immunity; autoimmunity, and gastrointestinal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of
25 the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., gastrointestinal, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a
30 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in leukocytes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of hematopoietic disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in small intestines and leukocytes indicates that it is expressed by various hematopoietic cells, for example, in the peyer's patches of intestine as well as within the circulation itself. Thus, it may play a role in the proliferation; survival; differentiation; or activation of various hematopoietic cell lineages. This may affect the cells' ability to recognize antigen; mount an immune response; participate in inflammatory processes; and effectively patrol the body for infectious or foreign agents. Alternately, expression of this gene product in small intestine may reflect a role in digestion and food processing. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1384 of SEQ ID NO:70, b is an integer of 15 to 1398, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

The translation product of this gene shares sequence homology with the *Drosophila strabismus* gene product which is thought to regulate tissue polarity and cell fate decisions (See Genebank Accession No.gi|2854044 (AF044208); all

references available through this reference are hereby incorporated herein by reference). When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

Preferred polypeptides of the invention comprise the following amino acid sequence: MQSPLVECPFPSIHYWPSVPAGAQQACSPMFHAAGWSRSQPNGEIPASSXGHL SIQRAAL VVLENYKDF TIYNPNLLTASKFRAAKHMAGLKVINVDGPSNNATGQSRAMIAAAARRRDSSHNELYYE EAEHERRVKRKARLVVAVEEAFIHIQRLQAEQQKAPGEVMDPREAAQAIFPSMARALQKYLRI TRQQ NYHSMESILQAPGLLHHQRHDPQGLPRTVPQC GPHPAI (SEQ ID NO: 407), LSIQRAALV VLENYKDF TIYNP (SEQ ID NO: 408), DSSHNELYYEAEHE (SEQ ID NO: 409), and/or FPSMARALQKYLRI TRQQ (SEQ ID NO: 410). Polynucleotides encoding these polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MAFKLLILLIGTWALFFRKRRADMPRVFVFRALLLVLI FLFCGFP IGFFT GSAFWTLGNRNYQGIVQYAVSPCGMPSSFHPLLAIRPCWSSGSLQPNVPRCRLVPLPTEWGNPRFQXGT PEYPASSIGGPRKLLQRFHHL (SEQ ID NO: 411). Polynucleotides encoding these polypeptides are also provided.

The translation product of this gene was determined to have a transmembrane domain located at amino acid position 249 - 266 of the amino sequence referenced in Table 1 for this gene. Likewise, this protein is thought to be a Type II membrane protein.

This gene is expressed primarily in human osteoclast stromal cells, fetal liver and spleen, and in endometrial tumors and to a lesser extent in hematopoietic cells, including T-cells and CD34 positive cells isolated from cord blood, as well as the thymus, fetal heart, 8 week old whole embryos, and tumors of pancreatic and testicular origin.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders, including AIDS and other hematopoietic diseases and/or disorders, in addition to tumors of osteoclast, endometrial, pancreatic, or testicular origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system as well as biological processes involved in cellular proliferation and/or differentiation, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, haematopoietic, skeletal, cancerous, and/or other tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid, lymph, breast milk, and/or seminal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 190 as residues: Pro-17 to Gln-24, Asp-86 to Ser-96, Arg-106 to Asn-112, Ala-119 to Ala-130, Ala-148 to Pro-155, Gln-223 to Leu-230. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells and tissues, combined with the detected GAS biological activity, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the

natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue expression in liver tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue traumas. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1543 of SEQ ID NO:71, b is an integer of 15 to 1557, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 62

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MGLPVSWAPPALWVLGCCALLLSLWALCTACRSPRTL (SEQ ID NO: 412). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in human thymus, human synovial
10 sarcoma, and to a lesser extent in breast cancer cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune diseases and/or disorders, particularly autoimmune disorders
15 such as arthritis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune,
20 hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 191 as residues: Pro-40 to Arg-50, Ser-72 to Arg-77, His-82 to Leu-91, Gln-171 to Glu-189, Val-203 to Gly-222, Pro-263 to Thr-269, Ser-282 to Trp-287. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in thymus indicates polynucleotides and polypeptides
30 corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19,

20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in cancerous and/or proliferative cells and tissues. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1149 of SEQ ID NO:72, b is an integer of 15 to 1163, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

The translation product of this gene shares sequence homology with human, porcine, and mouse zona pellucida binding protein sp 38 which is known to be important in sperm binding to the zona pellucida of an egg cell. Monoclonal antibodies directed against this protein have resulted in inhibition of the sperm/egg binding reaction. As such The translation product of this gene may show commercial utility as a contraceptive. (See Genebank Accession No. gn|PID|d1005021; all references available through this accession are hereby incorporated by reference herein).

15 Preferred polypeptides of the invention comprise the following amino acid sequence: IYGKLGQPDKIYVELHQNSP (SEQ ID NO: 413), FLEPLSGLYTCTLSYK (SEQ ID NO: 414), LQVVRLDSCRPGFGKN (SEQ ID NO: 415), and/or CVSULTYGAKSC (SEQ ID NO: 416). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in a human testes library. It has not been found in other libraries screened at HGS.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, infertility, and/or other reproductive diseases and/or disorders.

25 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male and female reproductive systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., testes, and cancerous and wounded tissues) or bodily fluids (e.g. seminal fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

30

level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 192 as residues: Lys-35 to Asp-40, Pro-75 to Asn-84,
5 Lys-114 to Arg-129, Arg-138 to Ser-143, Ser-154 to Asn-160, Val-224 to Asn-231, Arg-238 to Asp-243, Asp-276 to Asn-291, Lys-324 to Asp-338. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in testes combined with the homology to the human, porcine, and mouse zona pellucida protein Sp 38 indicates that polynucleotides and
10 polypeptides corresponding to this gene are useful for the production of a contraceptive vaccine. Alternatively, the protein may show utility in the diagnosis, treatment, and/or prevention of a variety of reproductive disorders within both the male and female reproductive systems. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male
15 contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as
20 hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may
25 show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1472 of SEQ ID NO:73, b is an integer of 15 to 1486, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid, and to a lesser extent, other cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

15

This gene is expressed primarily in an apoptotic T-cell library, and to a lesser extent, in whole embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, and developmental diseases and/or disorders, particularly disorders related to aberrant cell death regulation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, developmental, reproductive, apoptotic cells, and cancerous and healing tissue or cells) or bodily fluids (e.g., serum, lymph, amniotic fluid, plasma, urine, synovial fluid and spinal fluid, and/or lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20

25

30

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 193 as residues: Met-1 to Ala-6, Gly-51 to Gly-71. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in apoptotic T-cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the

protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1539 of SEQ ID NO:74, b is an integer of 15 to 1553, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

The translation product of this gene shares sequence homology with a 50 kDa glycoprotein of the human erythrocyte membrane associated blood-group antigen which is thought to have a transport or channel function in the erythrocyte membrane (See GenBank No. gb|X64594|HSEPMG50; all references available through this accession are hereby incorporated herein by reference). When tested against Jurkat cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates T-cells, and to a lesser extent, other cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. The translation product of this gene has been determined to contain two transmembrane domains located at amino acid positions 95 - 124, and 1 - 27 of the amino acid sequence referenced in Table 1 for this gene. Therefore, this protein may share structural characteristics to Type IIIa membrane protein. Based on the sequence similarity to the human erythrocyte membrane

associated blood-group antigen, and the structural similarity to type IIIa membrane proteins, The translation product of this gene is expected to share at least some biological activities with such proteins. Such activities are known in the art, some of which are described elsewhere herein.

- 5 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PAKGEGCRRLLHDHPHIWRLLWAHSDPDPLPTQPRAEQGETEFCVPVGPLCH
10 DWHPLPVDVLAQLQLSHILPWGQPAPSRHQHLLLLGSLRAYLGGNIQCPAKKGKLDMVHIQNATLAGGV
AVGTAAEMMLMPYGALIIGFVCGIISTLGFVYLTPFLESRLHIQDTCGINNLHGIPGIIGGIVGAVTAA
SASLEVYGKEGLVHSFDFQGFNGDWTARTQGKFQIYGLLVTLAMALMGGIIVGLILRLPFWGQPSDENC
FEDAVYWEMPEGNSTVYIPEDPTFKPSGPSVPSVPMVSPLMASSVPLVP (SEQ ID NO: 417).

Polynucleotides encoding these polypeptides are also provided.

- 15 The gene encoding the disclosed cDNA is believed to reside on chromosome 18. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 18.

This gene is expressed primarily in in tonsils and to a lesser extent in the larynx, kidney medulla, epithelial cells, keratinocytes, and cells involved in
20 hematopoiesis, especially neutrophils.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic diseases and/or disorders, in addition to, the
25 proliferation and/or differentiation of integumentary cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain
30 tissues or cell types (e.g., haematopoietic, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, lymph) or another tissue or cell sample taken from an individual having such a

disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 194 as residues: Gly-85 to Lys-94, Gln-125 to Cys-131, Glu-151 to Gly-159. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in tonsils, combined with the homology to a 50 kDa glycoprotein of the human erythrocyte membrane protein indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:75 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1636 of SEQ ID NO:75, b is an integer of 15 to 1650, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 66

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

10 PRVRTRAPVVPAGHRALSPAGVLLAVPAMLSLDFLDDVRRMNKRQVSLS
VLFFSWLFLSLRGCCGARRTPGFWCEGLSWSDTRVIRFLWRLWPEAALSASLFLTPN (SEQ ID
NO: 418). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in hematopoietic tissues, especially helper T-cells and anergic T-cells.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tuberculosis, AIDS, and other immune diseases and/or disorders, particularly infections and/or malignancies. Similarly, polypeptides and antibodies
20 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., haematopoeitic, immune, and cancerous, and/or wounded tissues) or
25 bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, and/or lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
30 epitopes shown in SEQ ID NO: 195 as residues: Asp-9 to Gln-17. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells and tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2136 of SEQ ID NO:76, b is an integer of 15 to 2150, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 15 - 34 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 1 - 14 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type II membrane proteins.

This gene is expressed primarily in the fetal liver/spleen, human brain, and retina.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, neurologic, and visual diseases and/or disorders, particularly retinoblastoma as well as other diseases or disorders involving the retina and/or brain. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurologic system and in eye development, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, visual, retinal, neural, cancerous, and/or wounded tissues) or bodily fluids (e.g., serum, plasma, aqueous humor, vitreous humor, urine, amniotic fluid, synovial fluid and spinal fluid, vitreous and aqueous humors) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 196 as residues: Glu-48 to Thr-54. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in fetal liver/spleen indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, expression of this gene with in the retina may suggest gene is useful for the diagnosis, treatment, and/or prevention of a variety of eye disorders and/or conditions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

immunotherapy targets for the above listed tissues. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies
5 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of
10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1578 of SEQ ID NO:77, b is an
15 integer of 15 to 1592, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

The translation product of this gene shares sequence homology with the
20 glutamate-binding subunit of an N-methyl-D-aspartate receptor complex. The amino acids L-glutamic and L-aspartic acids form the most widespread excitatory transmitter network in mammalian brain. The excitation produced by L-glutamic acid is important in the early development of the nervous system, synaptic plasticity and memory formation, seizures and neuronal degeneration. The receptors activated by L-
25 glutamic acid are a target for therapeutic intervention in neurodegenerative diseases, brain ischaemia and epilepsy. As such, the protein product of this gene may also play a role in the regulation of the nitrous oxide synthase gene which is known to be a vital link in various signal transduction pathways within the brain as well as other tissues (See GenBank No. bbs|61979 and Medline Article No.92049755). Moreover, The
30 translation product of this gene was also shown to have homology to a neural membrane protein 35 (See Genbank Accession No. gb|AAC32463.1| (AF044201); all references available through this accession are hereby incorporated herein by

reference; for example, Mol. Cell. Neurosci. 11 (5), 260-273 (1998)). The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 42 - 73, and 75 - 94 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to IIIa membrane proteins. When tested against U937 and Jurkat cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid and T-cells, and to a lesser extent, other cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

Preferred polypeptides of the invention comprise the following amino acid sequence: HASAWNLLILLTVFTLS (SEQ ID NO: 419), VYAALGAGVFTLFLALDTQLLMGN (SEQ ID NO: 420), EEYIFGALNIYLDIIYIF (SEQ ID NO: 421), and/or WNLILLTVFTLSMAYLTGMLSSYYNT (SEQ ID NO: 422). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

MAYLTGMLSSYYNTTSVLLCLGITALVCLSVTVFSFQTKFDFTSCQGVLF
VLLMTLFFSGLILAILLPFYVPWLHAVYAALGAGVFTLFLALDTQLLMGNRRHSLSP E EYIFGALNIY
LDIIYIFTFFLQLFGTNRE (SEQ ID NO: 242). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in the brain and to a lesser extent in dendritic cells and in the kidney cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, schizophrenia, epilepsy, brain ischaemia, and neurodegenerative diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g. neural, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 197 as residues: Ala-12 to Glu-27, Pro-35 to Ser-43, Pro-70 to Gly-79, Ser-92 to Val-98, Pro-166 to Leu-175, Ser-234 to Thr-246.

Polynucleotides encoding said polypeptides are also provided.

The tissue distribution combined with the homology to a known N-methyl-D-aspartate receptor indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. This protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment,

and/or prevention of developmental diseases and disorders. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation.

Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents
5 that modulate their interactions, in addition to its use as a nutritional supplement.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
10 related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1565 of SEQ ID NO:78; b is an integer of 15 to 1579, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

20 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 37 - 62 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to Type Ia membrane proteins. The translation product of this gene was also determined to have a conserved peroxidase-I
25 domain located at about amino acid position 15 - 25 of the amino acid sequence referenced in Table 1 for this gene.

Preferred polypeptides of the invention comprise the following amino acid sequence: TLSLLVSLHTV (SEQ ID NO: 423). Polynucleotides encoding these polypeptides are also provided.

30 This gene is expressed primarily in the brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological diseases and disorders, a non-limiting example of which includes, epilepsy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
5 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous, and/or wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
10 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or
15 prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease,
20 Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including
25 disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine
30 biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may

show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1382 of SEQ ID NO:79, b is an integer of 15 to 1396, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

When tested against Jurkat cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates T-cells, and to a lesser extent, other cells and tissue cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. Additional embodiments of the invention include polypeptides comprising the following amino acid sequences:

MSSSGTSDASPSGSPVLASYKPAPPKDKLPETPRRRMKKSLAPLHPEFEEVYRFGAESRKLLREPVD
AMPDPTPFLLARESAEVHLIKERPLVIPPIASDRSGEQHSPAREKPHKAHVGVVAHRIHHATPPQPARG
DPGGRPGERRQGGEALRDGQNCVKPAVPHPALSMHCEHHWEISATPFLFNPMAKHFSHLPTHSPSAS
LALFFTPKYDRVPAAEYVFPNCCGQTPVCRIACF (SEQ ID NO: 424); MSSSGTSDASPSGSPV
LASYKPAPPKDKLPETPRRRMKKSLAPLHPEFEEVYRFGAESRKLLREPVDAMPDPTPFLLARESAE
(SEQ ID NO: 425); VHLIKERPLVIPPIASDRSGEQHSPAREKPHKAHVGVVAHRIHHATPPQPARG
GEDPGGRPGERR (SEQ ID NO: 426); QGGEALRDGQNCVKPAVPHPALSMHCEHHWEISAT
PFLFNPMAKHFSHLPTHSPSASLALFFTPKYDRVPAAEYVFPNCCGQTPVCRIACF (SEQ ID NO:

427); KRASQPPCTRNLKRSTDSGQRAGNSFCGNQWMLCPTPPHFCWLGSPPRSTSSKRGPS
(SEQ ID NO: 428); and PPSPTAASSTARPAKSRTPTSGWHIGSTTPRRSQPEVKTLAV
DQVNGGKVVRKHSGTDRTV (SEQ ID NO: 429). Additional embodiments are directed
to polynucleotides encoding these polypeptides.

5 The gene encoding the disclosed cDNA is believed to reside on chromosome
12. Accordingly, polynucleotides related to this invention are useful as a marker in
linkage analysis for chromosome 12.

 This gene is expressed primarily in Endometrial Tumor, fetal liver,
Hypothalamus, Larynx carcinoma III, Prostate Cancer.

10 Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, endometrial tumor, larynx carcinoma III, prostate cancer, in addition to
other proliferative diseases and/or disorders. Similarly, polypeptides and antibodies
15 directed to these polypeptides are useful in providing immunological probes for
differential identification of the tissue(s) or cell type(s). For a number of disorders of
the above tissues or cells, particularly of the reproductive, hepatic, and pulmonary
systems, expression of this gene at significantly higher or lower levels is routinely
detected in certain tissues or cell types (e.g., hepatic, developmental, differentiating,
20 proliferative, and cancerous, and/or other tissues) or bodily fluids (e.g., serum,
plasma, urine, synovial fluid and spinal fluid, pulmonary surfactant) or another tissue
or cell sample taken from an individual having such a disorder, relative to the
standard gene expression level, i.e., the expression level in healthy tissue or bodily
fluid from an individual not having the disorder.

25 Preferred polypeptides of the present invention comprise immunogenic
epitopes shown in SEQ ID NO: 199 as residues: Ala-62 to Tyr-71. Polynucleotides
encoding said polypeptides are also provided.

 The tissue distribution in tumors of endometrium, larynx, and prostate origins,
combined with the detected GAS biological activity, indicates that polynucleotides
30 and polypeptides corresponding to this gene are useful for diagnosis and intervention
of these tumors, in addition to other tumors where expression has been indicated. The
expression within cellular sources marked by proliferating cells indicates this protein

may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and
5 elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Alternatively, the tissue distribution within liver tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and
10 conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate
15 cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
20 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
25 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1216 of SEQ ID NO:80, b is an integer of 15 to 1230, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

30 **FEATURES OF PROTEIN ENCODED BY GENE NO: 71**

In another embodiment, polypeptides of the invention comprise the following amino acid sequence: MWNPNAGQPGPNPYPPNIGCPGGSNPAHPPPINPPFPFGPCPPPPGAPHGN

PAFPPGGPPHPVPQPGYPGCQPLGYPYPPYPPAPGIPPVNPLAPGMVGPVIVDKKMQKKMKKAHKKM
 HKHQKHKKYHKHKGKSSSSSSSSSSSDSD (SEQ ID NO: 430); RVGPDAWADAEQAQAVERLE
 DTPKHVESQCRAARAKSISPQYWVPWRFQSCPPTY (SEQ ID NO: 431); STLSPRPLSSSPR
 SSPWQSSFPFRWAPSSCATARVSRMPTVGSLPSSIPTACPWNPSCESLGSWHGWTSDDSRQEDAENEE
 5 SS (SEQ ID NO: 432); MPGSQGQIHIPPILGALEVPILPTHLLIHPFPQAPVLLPQELPMA
 IQLSPQVGPLILCHSQGIQDANRWVPTLLHTRLPLESL (SEQ ID NO: 433); and/or
 MASIPPLPPPLPAVILTEYRPWTLPSSLTSSALPSSFRCHVVLGECSPCAPHPLPXPEHPAVEP
 (SEQ ID NO: 434). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in bone marrow and primary dendritic cells,
 10 in addition to macrophages.

Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of immune and haematopoietic diseases and/or
 disorders. Similarly, polypeptides and antibodies directed to these polypeptides are
 15 useful in providing immunological probes for differential identification of the
 tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
 particularly of the immune, expression of this gene at significantly higher or lower
 levels is routinely detected in certain tissues or cell types (e.g., haematopoietic,
 immune, and cancerous, and/or other tissues) or bodily fluids (e.g., serum, plasma,
 20 urine, synovial fluid and spinal fluid, and/or lymph) or another tissue or cell sample
 taken from an individual having such a disorder, relative to the standard gene
 expression level, i.e., the expression level in healthy tissue or bodily fluid from an
 individual not having the disorder.

The tissue distribution in bone marrow indicates polynucleotides and
 25 polypeptides corresponding to this gene are useful for the treatment and diagnosis of
 hematopoietic related disorders such as anemia, pancytopenia, leukopenia,
 thrombocytopenia or leukemia since stromal cells are important in the production of
 cells of hematopoietic lineages. Representative uses are described in the "Immune
 Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19,
 30 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo
 culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or
 chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis,
 therefore, it can be used in immune disorders such as infection, inflammation, allergy,

immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1125 of SEQ ID NO:81, b is an integer of 15 to 1139, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 72

In another embodiment, polypeptides of the invention comprise the following amino acid sequence:

PRHTYWGIVLVPAAAMASPHSHPAQGVLPQPGPQPRWEDRVALGTRGRSPGAYLTESAPQQASTTPGPPT
CHGKVGSEAWLGAAPGPLPHTPSHYAIRVPSNICSCPGASSAPALRGVVRQPPGPQNPRQGGRRGTRA
25 SPVGSFLFCV (SEQ ID NO: 435); MFAVLPAVEGRATPHQDRTCYPSSRPWPSQPSPRGSM
PVPRPGAARGQLDGHVQGGWALQWGGPPAPAVYRRMALPPRAAGSYLDRKCPHPLPGARLCPGLPL
(SEQ ID NO: 436); VFGAVFLTTPSHDLATPTGASGWCLLPWPAPTTLHRGSCSPQAHSLVG
RTGWPGQEGGAQGLTSLRVLPSRHPLPQGPPHVMARLVNPGWGEQPLAHCPPHTLTMQFEFQATFAP
ALGPALPQP (SEQ ID NO: 437); HEEPPAGFGLRLSLWRRSPHEVGARLPNGAFGFSVRCLLCF
30 PPWRAEPPHIRIGRATPPGPGPGPASPALPEARCLCQGGQPEGSWMATCRVKAGPCSGAGRQPPQFTDA
WLFLPEQPAATWTGNVLIPSLGPGSALAFLCEPLLSLCCLGTPDRGVRVCPSVTFYSPRVEERKRKSK
GVQTPPQ (SEQ ID NO: 438); MATCRVKAGPCSGAGRQPPQFTDAWLFLPEQPAATWTGNVLIP
SLGPGSALAFLCEPLLSLCCLGTPDRGVRVCPSVTFYSPRVEERKRKSKGVQTPPQ (SEQ ID NO:
439); MKWFSTQPLWLNTRKQSRHRRGPGPPAPLSGVLGSRGLPHHPSQGWGRAGPRAGANVAWNSN

CIVRWVGGQWARGCSQPGPFTTNLAMTCGGPWGSGCLLGSTLSEVSPWAPPSCPQGHVLPTRLWAWGL
 QDPLCRVRVVGAGHGSRHQPDAPVGVARSWDGVVRNTAPKTQNKNTTNGRRSPPPTEVGFPELLIFPVSF
 LQPLVSRKSQTGTHAHHGQESRDSTKKGGVHRGRPGQSLAPGRG (SEQ ID NO: 440); KVTDGH
 TRTPRSGVPRQHKERRGSQRKARAEPPREGMRTFPVQVAAGCSGRKSHASVNCWGWRPAPLQGPALTL
 5 HVAIQLPSCGPWPWHRHRASRAGLAGPGPGPGGVARPILMWGGSALHGGKHSKHRTLKPKAPLGLSLAPT
 SWGGDRRHRDLSPKPAGGSSC (SEQ ID NO: 441); and/or MRTFPVQVAAGCSGRKSHASV
 NCWGWRPAPLQGPALTLHVAIQLPSCGPWPWHRHRASRAGLAGPGPGPGGVARPILMWGGSALHGGKHS
 KHRTLKPKAPLGLSLAPTSWGGDRRHRDLSPKPAGGSSC (SEQ ID NO: 442).

Polynucleotides encoding these polypeptides are also provided.

10 The gene encoding the disclosed cDNA is believed to reside on chromosome
 7. Accordingly, polynucleotides related to this invention are useful as a marker in
 linkage analysis for chromosome 7.

This gene is expressed primarily in healing wound tissues, macrophage-
 oxLDL, hemangiopericytoma, and CD34+ cells.

15 Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, healing wound, and proliferative diseases and/or disorders, particularly
 soft tissue cancers, such as hemangiopericytoma. Similarly, polypeptides and
 20 antibodies directed to these polypeptides are useful in providing immunological
 probes for differential identification of the tissue(s) or cell type(s). For a number of
 disorders of the above tissues or cells, particularly of healing wounds, expression of
 this gene at significantly higher or lower levels is routinely detected in certain tissues
 or cell types (e.g., lymph, cancerous, and/or wounded tissues) or bodily fluids (e.g.,
 25 serum, plasma, urine, synovial fluid and spinal fluid, and/or lymph) or another tissue
 or cell sample taken from an individual having such a disorder, relative to the
 standard gene expression level, i.e., the expression level in healthy tissue or bodily
 fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic
 30 epitopes shown in SEQ ID NO: 201 as residues: Met-1 to Gly-6, Arg-23 to Gly-33,
 Arg-60 to Ala-66, Thr-90 to Gly-103, Glu-105 to Trp-112. Polynucleotides encoding
 said polypeptides are also provided.

The tissue distribution within healing wounds indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Representative uses are described elsewhere herein. Expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:82 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1395 of SEQ ID NO:82, b is an integer of 15 to 1409, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:82, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 73

The translation product of this gene has homology to the Pro-Pol-dUTPase polyprotein of a newly discovered retrovirus. Since this protein also shares homology

to the human HERV-L element, and considering that most retroviruses integrate their proviral form into eukaryotic genomes through a homologous recombination mechanism, this gene is useful in providing protection against retroviral infections or could be used in the development of gene therapy applications (See Genebank

- 5 Accession No.2065210; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: GLMECLIHRHGS (SEQ ID NO: 443), and/or STKGMQFILTGITLSGY (SEQ ID NO: 444). Polynucleotides encoding these polypeptides are also provided.

- 10 This gene is expressed primarily in CD34 positive cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune diseases and/or disorders, particularly viral infections.

- 15 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, and cancerous, wounded, and/or other tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid, and/or lymph) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 25 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 202 as residues: Arg-39 to Thr-49, Leu-52 to Gly-60, Ser-67 to Arg-76, Gln-130 to Phe-137, Ser-139 to His-148. Polynucleotides encoding said polypeptides are also provided.

- 30 The tissue distribution in CD34+ immune cells combined with the homology to a retroviral protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in immune indicates a role in the regulation

of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g. by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:83 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 700 of SEQ ID NO:83, b is an integer of 15 to 714, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:83, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

The translation product of this gene shares sequence homology with mouse, bovine, and human butyrophilins, which are thought to be important in lactation especially during the latter part of pregnancy. Butyrophilin is a glycoprotein of the immunoglobulin superfamily that is secreted in association with the milk-fat-globule

membrane from mammary epithelial cells (See Genbank Accession No.gb|AAB51034.1, and Geneseq Accession No. W97814; all references available through these accessions are hereby incorporated herein by reference; for example, Mamm. Genome 7 (12), 900-905 (1996)). Based on the sequence similarity, The
5 translation product of this gene is expected to share at least some biological activities with glycoproteins. Such activities are known in the art, some of which are described elsewhere herein.

In another embodiment, polypeptides of the invention comprise the following amino acid sequence: PRVRALLFARSLRLCRWGAKRLGVASTEAQRGVSFKLEEKTAHSSLALFRD
10 DTGVKYGLVGLEPTKVALNVERFREWAVVLADTAVTSGRHYWEVTVKRSQQFRIGVADVDMSRDSCIGV
DDRSWVFTMPSASGTPCWPTRKPQLRVLGSQEVGLLLEYEAQKLSLVDVSQVSVVHTLQTDFRGPVVPFA
FALWDGELLTHSGLEVPEGL (SEQ ID NO: 445), and/or MSRDSCIGVDDRSWVFTMPSASG
TPCWPTRKPQLRVLGSQEVGLLLEYEAQKLSLVDVSQVSVVHTLQTDFRGPVVPFAFALWDGELLTHSGL
EVPEGL (SEQ ID NO: 446). Polynucleotides encoding these polypeptides are also
15 provided.

This gene is expressed primarily in adult heart, LNCAP cell line, OB cell line (HOS fraction), and epididymis, and to a lesser extent in a variety of other cells and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, coronary disease and heart tumors and reproductive disorders, particularly those of the male reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
25 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly those of the heart and reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cardiovascular, cardiac, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine,
30 seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 203 as residues: Gly-30 to Ser-36. Polynucleotides
5 encoding said polypeptides are also provided.

The tissue distribution and homology to butyrophilin indicates that polynucleotides and polypeptides corresponding to this gene are useful for determining the mechanisms underlying mammary-specific gene expression, lactation, and potentially for the production of copious amounts of butyrophilin or
10 heterologous proteins in the milk of transgenic animals. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding
15 Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy);
20 regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction
25 etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to
30 isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies

directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:84 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1083 of SEQ ID NO:84, b is an integer of 15 to 1097, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:84, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

The translation product of this gene shares sequence homology with angiopoietin-2 which is thought to be important in regulation of angiogenesis through the Tie2, or other receptor tyrosine kinase (See Genbank Accession Nos. gb|AAC97965.1| (AF110520), and gb|AAB63189.1| (AF004326); in addition to Geneseq Accession No. R94603; all references available through these accessions are hereby incorporated herein by reference; for example, Science 277 (5322), 55-60 (1997)). Based on the sequence similarity, The translation product of this gene is expected to share at least some biological activities with angiogenic and kinase proteins. Such activities are known in the art, some of which are described elsewhere herein.

In another embodiment, polynucleotides of the invention comprise the following nucleic acid sequence:

```
GCACGAGCGGCACGAGCGGATCCTCACACGACTGTGATCCGATTCTTTCCAGCGGCTTCTGCAACCAAG
CGGGTCTTACCCCCGGTCTCCGCGTCTCCAGTCTCGCACCTGGAACCCCAACGTCCCCGAGAGTCCC
CGAATCCCCGCTCCAGGCTACCTAAGAGGATGAGCGGTCTCCGACGGCCGGGGCAGCCCTGATGCTC
TGCGCCGCCACCGCCGTGCTACTGAGCGCTCAGGGCGGACCCGTGCAGTCCAAGTCGCCGCGCTTTGCG
TCCTGGGACGAGATGAATGTCTCGCGCACGGACTCTGCAGCTCGGCCAGGGGCTGCGCGAACACGCG
GAGCGCACCCGAGTCAGCTGAGCGCGCTGGAGCGGCGCTGAGCGCGTGGGGTCCGCCTGTTCAGGGA
ACCGAGGGGTCCACCGACCTCCCGTTAGCCCTGAGAGCCGGGTGGACCTGAGGTCCTTCACAGCCTG
```

CAGACACAACCTCAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCCAGCAGCAGCGG
CACCTGGAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGACCACAAGCAC
CTAGACCATGAGGTGGCCAAGCCTGCCCCAAGAAAGAGGCTGCCCGAGATGGCCCAGCCAGTTGACCCG
GCTCACAATGTCTAGCCGCTGCACCGGCTGCCCAGGGATTGCCAGGAGCTGTTCCAGGTTGGGGAGAGG
5 CAGAGTGGACTATTTGAAATCCAGCCTCAGGGGTCTCCGCCATTTTGGTGAACCTGCAAGATGACCTCA
GATGGAGGCTGGACAGTAATTCAGAGGCGCCACGATGGCTCAGTGGACTTCAACCGGCCCTGGGAAGCC
TACAAGGCGGGGTTTGGGGATCCCCACGGCGAGTTCTGGCTGGGTCTGGAGAAGGTGCATAGCATCACG
GGGGACCGCAACAGCCGCTGGCCGTGCAGCTGCGGGACTGGGATGGCAACGCCGAGTTGCTGCAGTTC
TCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCAGCTCACTGCACCCGTGGCCGGCCAGCTG
10 GCGGCCACCAACCGTCCCCACCCAGCGGCCCTCTCCGTACCCTTCTCCACTTGGGACCAGGATCACGACCTC
CGCAGGGACAAGAACTGCGCCAAGAGCCTCTCTGGAGGCTGGTGGTTTGGCACCTGCAGCCATTCCAAC
CTCAACGGCCAGTACTTCCGCTCCATCCACAGCAGCGGCAGAAAGCTTAAGAAGGGAATCTTCTGGAAG
ACCTGGCGGGGGCCGCTACTACCCGCTGCAGGCCACCACTGTTGATCCAGCCCATGGCAGCAGAGGCA
GCCCTCTAGCGTCTGGCTGGGCCTGGTCCCAGGCCACGAAAGACGGTGACTCTTGGCTCTGCCCCGAG
15 GATGTGGCCGTTCCCTGCCTGGGCAGGGGCTCCAAGGAGGGGCCATCTGGAAACTTGTGGACAGAGAAG
AAGACCACGACTGGAGAAGCCCCCTTTCTGAGTGCAGGGGGGCTGCATGCGTTGCCTCCTGAGATCGAG
GCTGCAGGATATGCTCAGACTCTAGAGGCGTGGACCAAGGGGCATGGAGCTTCACTCCTTGTGCGCCAG
GGAGTTGGGGACTCAGAGGGACCCTTGGGGCCAGCCAGACTGGCCTCAATGGCGGACTCAGTCACATT
GACTGACGGGGACCAGGGCTTGTGTGGGTGCGAGAGCGCCCTCATGGTGCTGGTGTGTTGTGTGTAGGT
20 CCCCTGGGGACACAAGCAGGCGCCAATGGTATCTGGGCGGAGCTCACAGAGTTCTTGGGAATAAAGCAA
CCTCAGAACAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA (SEQ ID NO: 447),
and/or
ATGAGCGGTGCTCCGACGGCCGGGGCAGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCT
CAGGGCGGACCCGTGCAGTCCAAGTCGCCGCGCTTTGCGTCTCTGGGACGAGATGAATGTCTTGGCGCAC
25 GGACTCTCTGCAGCTCGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGAGTCAGCTGAGCGCGCTG
GAGCGGCGCCTGAGCGCTGCGGGTCCGCCTGTCTAGGGAACCGAGGGGTCCACCGACCTCCCGTTAGCC
CCTGAGAGCCGGGTGGACCTGAGGTCTTACAGCCTGCAGACACAACCTCAAGGCTCAGAACAGCAGG
ATCCAGCAACTCTTCCACAAGGTGGCCCAGCAGCAGCGGCACCTGGAGAAGCAGCACCTGCGAATTCAG
CATCTGCAAAGCCAGTTTGGCCTCCTGGACCACAAGCACCTAGACCATGAGGTGGCCAAGCCTGCCCGA
30 AGAAAGAGGCTGCCCAGATGGCCCAGCCAGTTGACCCGGCTCACAATGTCAGCCGCTGCACCGGCTG
CCCAGGGATTGCCAGGAGCTGTTCCAGGTTGGGGAGAGGCAGAGTGGACTATTTGAAATCCAGCCTCAG
GGGTCTCCGCCATTTTGGTGAACCTGCAAGATGACCTCAGATGGAGGCTGGACAGTAATTCAGAGGCGC
CACGATGGCTCAGTGGACTTCAACCGGCCCTGGGAAGCCTACAAGGCGGGGTTTGGGGATCCCCACGGC
GAGTTCTGGCTGGGTCTGGAGAAGGTGCATAGCATCACGGGGACCGCAACAGCCGCTGGCCGTGCAG
35 CTGCGGGACTGGGATGGCAACGCCGAGTTGCTGCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCC
TATAGCCTGCAGCTCACTGCACCCGTGGCCGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTC
TCCGTACCCTTCTCCACTTGGGACCAGGATCACGACCTCCGCGAGGACAAGAACTGCGCCAAGAGCCTC
TCTGGAGGCTGGTGGTTTGGCACCTGCAGCCATTCCAACCTCAACGGCCAGTACTTCCGCTCCATCCCA
CAGCAGCGGCAGAAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGCGGCTACTACCCGCTGCAG
40 GCCACCACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCCTAG (SEQ ID NO: 448).

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MAQWTSTGPGKPTRRGLGIPTASSGWVWRRCIASWGTATAAWPCSCGTGMA TPSCSSPCTWVARTRPIACSSLHPWPASWAPPPSHPAASPYPSPLGTRITTSAGTRTAPRASLEAGGL APAAIPTFNGPVLPAAPSHSSGRSLRRESSGRPAGRYYPLQATTMLIQPMAAEAAAS (SEQ ID NO:

5 449) . Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in osseous tissues, kidney cortex, bone
10 marrow, larynx carcinoma, and pineal gland, and to a lesser extent in placenta, stromal cells, epithelioid sarcoma, and a variety of other cells and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, arthritis, kidney and urinary tract disorders, immune cell and system dysfunctions, disorders of the pineal gland and brain, and carcinomas, particularly of the larynx. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
20 particularly those of the immune, connective, endocrine, and urinary systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
25 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 204 as residues: Pro-27 to Arg-34, Glu-60 to Gln-65, Cys-80 to Thr-87, Leu-109 to Ile-116, Ala-124 to Gln-133, Lys-158 to Leu-165, Arg-
30 229 to Ser-234, Asp-236 to Trp-241, Thr-266 to Ser-271, Thr-328 to Lys-343, Ser-355 to Tyr-363, Ile-367 to Lys-376, Thr-382 to Tyr-387. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to angiopoietin-2 indicates that polynucleotides and polypeptides corresponding to this gene are useful for the regulation of angiogenesis, particularly since angiogenesis is thought to depend on a precise balance of positive and negative regulation. Angiopoietin-1 (Ang1) is an angiogenic factor that signals through the endothelial cell-specific Tie2 receptor tyrosine kinase and, like vascular endothelial growth factor, is essential for normal vascular development in the mouse. Angiopoietin-2 is a naturally occurring antagonist for Angiopoietin-1 and Tie2. Transgenic overexpression of Angiopoietin-2 disrupts blood vessel formation in the mouse embryo. In adult mice and humans, Angiopoietin-2 is expressed only at sites of vascular remodeling. As such, this gene, or antagonists thereof, are useful in the diagnosis and treatment of arthritis, bone growth and remodeling, cancers (particularly those of bone, connective, lymphatic, and vascular tissues), ischaemia, lymphangiogenesis, lymphadenitis, lymphadenoma, lymphadenosis, lymphangitis, lymphangioendothelioma, lymphangioma, lymphangiophlebitis, lymphangiosarcoma, lymphatitis, lymphedema, lymphenteritis, angioma, angiomegaly, angiomyosarcoma, angiomyoma, angiomyolipoma, angiomyoneuroma, angioneuromyoma, angiosarcoma, angiostenosis, angioelectasis, and as a lymphagogue. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:85 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1917 of SEQ ID NO:85, b is an

integer of 15 to 1931, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:85, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 76

5 The translation product of this gene was shown to have homology to the DPM2 mannosyl transferase gene, which is known to be important in O-linked oligosaccharide glycosylation of proteins. Mutations within this gen have been shown to result in reduced levels of O-glycosylation. Since defects in proper protein glycosylation can result in the development of antigen-specific antibodies to such
10 protein or altered pharmacokinetics (i.e., plasma half-life, in vivo clearance rate, etc.), the protein product of this gene may show utility in the treatment, diagnosis, and/or prevention of various abnormalities involving oligosaccharide metabolism, specifically those associated with O-glycosylation (See Genebank Accession No.R47201).

 Preferred polypeptides of the invention comprise the following amino acid
15 sequence: GHDLPODAWLRLWVLGALCAGGWAVNYLPFFL (SEQ ID NO: 450), and/or FLYHYLPALTFQILLLPV (SEQ ID NO: 451). Polynucleotides encoding these polypeptides are also provided.

 The gene encoding the disclosed cDNA is believed to reside on chromosome 9. Accordingly, polynucleotides related to this invention are useful as a marker in
20 linkage analysis for chromosome 9.

 This gene is expressed primarily in brain and melanocytes and to a lesser extent in breast, testis, and colon.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, particularly of the brain and melanocyte, in addition to neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
30 tissues or cells, particularly of the brain, central nervous system, PNS, epithelial tissues including other parts of the integumentary system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types

(e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
5 having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 205 as residues: His-31 to Gln-38, Tyr-65 to Ser-71. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain tissue, combined with the homology to a
10 known enzyme involved in oligosaccharide metabolism, indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and
"Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and
15 elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia,
20 mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse
25 formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may
30 show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:86 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1078 of SEQ ID NO:86, b is an integer of 15 to 1092, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:86, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

Preferred polypeptides of the invention comprise the following amino acid sequence: DICRLERAVCRDEPSALARALTWRQARAQAGA (SEQ ID NO: 453), XAPATXAWDTVVPPLPRKCQCSGSARSHGAGRSALHSPLEGSRPKVPAGAVGKSLPGQSRPQHCLPPKQPKQCRPGL ELKEGPLLTPTRASVQLSHPACLYWAPLLWIRDPASV (SEQ ID NO: 454), XAPATXAWDTVV PPLPRKCQCSGSARSHGAGRSALHSPLEGSRPKVPAGAVGKSL (SEQ ID NO: 455), PGQSRPQHCLPPKQPKQCRPGL ELKEGPLLTPTRASVQLSHPACLYWAPLLWIRDPASV (SEQ ID NO: 456), and/or MSPLPWPGPLPGGRQGHRLPECCSSGCAGGPTWPHCSSQSWPMXSARHXGLGHC CPSSP (SEQ ID NO: 452). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence: DICRLERAVCRDEPSALARALTWRQARAQAGAMLLFGLCWGPYVATLLL SVLAYXQRPPLXPGTLLSLLSLGSASAAVPVAMGLGDQRYTAPWRAAAQRCLQGLWGRASRDSPGPSI AYHPSSQSSVDLDLN (SEQ ID NO: 457). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in cells of the immune system, including dendritic cells and T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders affecting the immune system, particularly immunodeficiencies such as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
5 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample
10 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in dendritic and T cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment
15 and/or prevention of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages,
20 including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g., by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for
25 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders,
30 such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted

factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:87 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 564 of SEQ ID NO:87, b is an integer of 15 to 578, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:87, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 78

Preferred polypeptides of the invention comprise the following amino acid sequence: MERVGMESGEMVCGLGSACNNPSDLGQVPVPLWXSVPVFGXGWNGH (SEQ ID NO: 458), MRSFQDVSALEEWGKGKDLPTHSLLLLLPLRDLLVVLGEIRKRQMEGCVWKGWGNPEK WFAVLALPVTRVTLGKSLSLSGXQFLHLYLERVGMGTEVLSSDLL (SEQ ID NO: 459), MHPAGPTFMGSKPIREQQFGPDACLLLLCVAMAGTEASRAAQQCTSQKVRAGQDFSASNPXQIQVEKL XPREGQGLAQGHSGCYRQSQDRKPFLRIPSPFPYTTLHLPFPDFAKNH (SEQ ID NO: 460), MHPAGPTFMGSKP IREQQFGPDACLLLLCVAMAGTEASRAAQQCTSQKVRAGQDFSASNP (SEQ ID NO: 461), PREGQGLAQGHSGCYRQSQDRKPFLRIPSPFPYTTLHLPFPDFAKNH (SEQ ID NO: 462), DPRVRKPPTATLTARTRPSTD (SEQ ID NO: 463), and/or AALEASVPAIATQRSSRQASGPNCCSLMGLDPMKVGPAGCISWDSVEADQVAGASGGRIEVKCGCMENL XRLHLGSGKGQXX (SEQ ID NO: 464). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in prostate and gall bladder.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the reproductive and gastrointestinal systems, including cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and urogenital systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 207 as residues: Arg-21 to Glu-30. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in gall bladder indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylketonuria, galactosemia, porphyrias, and Hurler's syndrome. In addition, expression of this gene product in the prostate - while likely to be reflective of non-specific expression of a variety of genes in the testes - may nevertheless be indicative of a role for this gene product in normal prostate function, and may implicate this gene product in male fertility, and could even suggest its use as a male contraceptive. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:88 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 685 of SEQ ID NO:88, b is an integer of 15 to 699, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:88, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 79

Preferred polypeptides of the invention comprise the following amino acid sequence: GXANPEDSVCILEGFSVTALSILQHLVCHSGAVRLPITVRSGGRFCCWGRKQEPGSQ XSDGD (SEQ ID NO: 466), AVQQQHRVPQTAHCPPLLVGWPWGSPCPPHCQPLSVQHHRERSDHL HITLAVGASDWQGALAHQA (SEQ ID NO: 467), PKTLPVISCPGSSVCSKCCQSASARHPC LACCWLLSSSPCWRTTTSWHLSSVPTQKAASCCCTCTSHHGLTEWPWRHNGSSWNKRWCWSLVLCK SPLPPVTGSNCQCNVEVVRALTVMHRQWLTVRRAGGPRTDQQRRTVRCLRDTVLLHGLSQDKLFM MHCVEVLHQFDQVMPGVSMILIRGLPDVTDCEEAALDDLCAAETDVEDPEVECG (SEQ ID NO: 468), and/or MLHRQWLTVRRAGGPRTDQQRRTVRCLRDTVLLHGLSQDKLFMMHCVEVL HQFDQVMPGVSMILIRGLPDVTDCEEAALDDLCAAETDVEDPEVECG (SEQ ID NO: 465). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence: GXANPEDSVCILEGFSVTALSILQHLVCHSGAVRLPITVRSGGRFCCWGRK QEPGSQXSDGDMTSALRGVADDQGHPLKMLLHLLAFSSAATGHLQASVLTQCLKVLVKLAENTSCDF LPRFQCVFQVLPKCLSPETPLPSVLLAVELLSLLADHDQLAPQLCSHSEGCLLLLLMYITSRPDRVAL ETQWLQLEQEVVWLLAKLGVQ EPLAPSHWLQLPV (SEQ ID NO: 469). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in breast, prostate, and to a lesser extent in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the reproductive organs of both males and females, especially cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, seminal fluid, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution primarily in breast, prostate, and to a lesser extent in testes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting the reproductive organs of males and females, including but not limited to cancers. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:89 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1112 of SEQ ID NO:89, b is an integer of 15 to 1126, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:89, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 80

The translation product of this gene shares sequence homology with epsilon-COP which is part of coatamers which are thought to be important in maintaining Golgi structure and in mediating ER-through- Golgi transport, and which can influence normal endocytic recycling of LDL receptors (See Genebank Accession No. gi|2443869 (AC002985); all references available through this accession are hereby
10 incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: MSGQLDARPAALHPQGLAHPLWTCLLPKGPSEVPQRPPQLWVVSISVLQGQHRGR AGPRDEQSVDVTNTTFLMAASIYLHDQNPDAALRALHQGDSLEW (SEQ ID NO: 470),
15 SVDVTNTTFLMAASIYLHD (SEQ ID NO: 471), QNPDAALRALHQGDSLE (SEQ ID NO: 472), and/or RDSIVAELDREMSR (SEQ ID NO: 473). Polynucleotides encoding these polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MLGLLLLCTPRAWLTLSGPVCFQGRDPLRSHRGHPSCGS (SEQ ID
20 NO: 474). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in breast tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, disorders affecting the immune and reproductive systems, particularly of the mammary glands. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and reproductive systems, expression of
30 this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or

cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 209 as residues: Gly-24 to Gln-36, Gly-47 to His-66. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in breast tissue and homology to epsilon-COP indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting the immune and reproductive systems, including cancers, which arise from abnormalities in coatomer function, particularly of those tissues actively involved in secretory functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:90 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1023 of SEQ ID NO:90, b is an integer of 15 to 1037, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:90, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

The translation product of this gene shares sequence homology with the highly conserved epoxide hydrolase which is thought to have an important function in the catalysis of potentially toxic or carcinogenic epoxides into their corresponding, inert

diols (See e.g., Genbank Accession No. gi|485136; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: HGFPEFWYSWR (SEQ ID NO: 475), ASHWLQQDQP (SEQ ID NO: 476),
5 PINHYRNIF (SEQ ID NO: 477), YPEMVMKLI (SEQ ID NO: 478),
PEFWYSWRYQLREF (SEQ ID NO: 479), HDWGGMIAW (SEQ ID NO: 480).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in benign and malignant prostate tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the prostate and liver, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
15 a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, prostate, cancerous and wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
20 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 210 as residues: Gln-38 to Pro-49, Glu-104 to Tyr-
25 109, His-127 to Lys-132, Thr-236 to Cys-243, Gln-328 to Asp-333, Lys-344 to Asp-351. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in tumors of prostate origins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of these tumors, in addition to other tumors where expression has
30 been indicated. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Alternatively, homology to epoxide hydrolase indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:91 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1302 of SEQ ID NO:91, b is an integer of 15 to 1316, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:91, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

This gene is expressed primarily in merkel cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample

taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 211 as residues: Lys-23 to Lys-29. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in immune tissue indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g. by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:92 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1007 of SEQ ID NO:92, b is an

integer of 15 to 1021, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:92, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 83

5 This gene is expressed primarily in liver tissue, particularly hepatomas.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the liver, including cancers. Similarly, polypeptides and
10 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic and hematopoietic systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, cancerous and wounded tissues)
15 or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred polypeptides of the present invention comprise immunogenic
20 epitopes shown in SEQ ID NO: 212 as residues: Met-1 to Ser-7, His-66 to Phe-72. Polynucleotides encoding said polypeptides are also provided.

 The tissue distribution in liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and
25 conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate
30 cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed

against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:93 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1246 of SEQ ID NO:93, b is an integer of 15 to 1260, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:93, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 84

Preferred polypeptides of the invention comprise the following amino acid sequence: GSLPPKPIYLVVPR (SEQ ID NO: 481). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in skin.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the skin, such as melanoma and wound healing, in addition to other disorders affecting the integumentary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and skin, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., epithelial, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 213 as residues: Cys-56 to Pro-73, Pro-83 to Lys-92. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in skin and skin melanoma indicates that

5 polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of various skin disorders including skin tumors, in addition to other tumors where expression has been indicated. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", "Infectious Disease", and "Regeneration" sections below, in Example 11, 19, and 20, and elsewhere herein.

10 Briefly, the protein is useful in detecting, treating, and/or preventing congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema,

15 photosensitivity, autoimmune disorders (i.e., lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e., cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis,

20 erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders (i.e., arthritis, trauma, tendonitis, chondromalacia and inflammation, etc.), autoimmune disorders (i.e., rheumatoid arthritis, lupus,

25 scleroderma, dermatomyositis, etc.), dwarfism, spinal deformation, joint abnormalities, and chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify

30 agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:94 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 976 of SEQ ID NO:94, b is an integer of 15 to 990, where both a and b correspond to the positions of nucleotide
10 residues shown in SEQ ID NO:94, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 85

When tested against kidney K562 cell lines, supernatants removed from cells containing this gene activated the interferon-sensitive responsive element (ISRE)
15 pathway. Thus, it is likely that this gene activates kidney or endothelial cells through the ISRE signal transduction pathway. ISRE is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of
20 the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. This gene maps to chromosome 10, and therefore, is used as a marker in linkage analysis for chromosome 10.

This gene is expressed primarily in placenta, and to a lesser extent in many other tissues or cells.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular disease including occlusion of vessels and arteries. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
30 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels is routinely detected in

certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in
5 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 214 as residues: His-58 to Gly-68, Thr-76 to Arg-81. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in placenta combined with the biological activity data
10 indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within highly vascularized tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression in placenta indicates that this protein may play a role in
15 the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in
20 pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
25 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:95 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1696 of SEQ ID NO:95, b is an integer of 15 to 1710, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:95, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 86

This gene is Apolipoprotein M (See, e.g., Genbank Accession No. gb|AAD18084.1|(AF129756) and gb|AAD11443.1|(AF118393); all references available through these accessions are hereby incorporated by reference herein). The protein components of human lipoproteins, apolipoproteins, allow the redistribution of cholesterol from the arterial wall to other tissues and exert beneficial effects on systems involved in the development of arterial lesions, like inflammation and hemostasis.

The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 6.

This gene is expressed primarily in fetal liver, fetal spleen, and to a lesser extent in adult liver, hepatocellular tumors, retina and testis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, proliferative disorders of the blood and tumors of the liver or disorders of lipid metabolism. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, metabolic, and hepatic systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., liver, hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g., bile, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 215 as residues: Glu-106 to Lys-120, Glu-136 to Tyr-141, Asn-148 to Pro-154. Polynucleotides encoding said polypeptides are also provided.

5 The tissue distribution of the gene product, ApoM, in fetal liver, and adult liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment and prevention of lipid metabolism disorders, including but not limited to, vascular disease, such as coronary artery disease, arteriosclerosis, and/or atherosclerosis. Additionally, The tissue distribution in fetal
10 liver and spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in fetal tissues indicates a role in
15 regulating the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer e.g., by boosting immune responses.

20 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and
25 in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
30 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Alternatively, expression within liver tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g.

hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:96 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 767 of SEQ ID NO:96, b is an integer of 15 to 781, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:96, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 87

This gene is expressed primarily in LPS treated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic or acute inflammatory disease, and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency, etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:97 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1099 of SEQ ID NO:97, b is an integer of 15 to 1113, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:97, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 88

The translation product of this gene shares sequence homology with prolylcarboxypeptidase which is thought to be important in the processing of

bioactive peptides like angiotensin and bradykinin (See Genbank Accession No. gb|AAA99891.1|; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides comprise the following amino acid sequence:

- 5 LVFAEHRYYGKSLPFG (SEQ ID NO: 482), EQALADFAEL (SEQ ID NO: 483),
 GGSYGGMLSAYLRMKYPH (SEQ ID NO: 484), NIIFSNGNLDPWAGG (SEQ ID NO:
 485), AMMDYPYPTDFLGPLPANPVKV (SEQ ID NO: 486), and/or FYTGNEG (SEQ
 ID NO: 487). Also preferred are the polynucleotides encoding these polypeptides.
 An additional preferred polypeptide fragment of the invention comprises the
 10 following amino acid sequence:
 MGSAPWAPVLLLLALGLRGLQAGARSGPRLPGALLPAASGPLQLRALRQQDL
 PSALPGVGQVLGPGRGALLLHWERGRRVGLRQQGLRRGLAAERGALLVFAEHRYYGKSLPFGAQSTQ
 RGHTELLTVEQALADFAELLRALRRDLGAQDAPAIAGGSYGGMLSAYLRMKYPHLVAGALAASAPVLS
 VAGLGDSNQFFRDVTADFEGQSPKCTQGVREAFRQIKDLFLQGAYDTRWEFGTCQPLSDEKDLTQLFM
 15 FARNAFTVLAMMDYPYPTDFLGPLPANPVKVGCDRLLEAQRITGLRALAGLVYNASGSEHCYDIYRLY
 HSCADPTGCGTGPDARAWDYQACTEINLTFASNNVTDMFDPDLPTDELQRQYCLDTWGVWPRPDWLLTS
 FWGGDLRAASNIIFSNGNLDPWAGGIRRNLSASVIAVTIQGGAHHLDRASHPEDPASVVEARKLEAT
 IIGEWVKAARREQQPALRGGPRLSL (SEQ ID NO: 488). Polynucleotides encoding these
 polypeptides are also provided.

- 20 This gene is expressed primarily in uterine cancer, testis, and to a lesser extent
 in lymph nodes, dendritic cells and HL60 cell line.

- Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 25 not limited to, uterine cancer, reproductive, and immune disorders. Similarly,
 polypeptides and antibodies directed to these polypeptides are useful in providing
 immunological probes for differential identification of the tissue(s) or cell type(s). For
 a number of disorders of the above tissues or cells, particularly of the reproductive
 system, expression of this gene at significantly higher or lower levels is routinely
 30 detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded
 tissues) or bodily fluids (e.g., amniotic fluid, seminal fluid, lymph, serum, plasma,
 urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
 individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 217 as residues: Gly-23 to Ala-30, Pro-44 to Phe-54,
5 Glu-69 to Pro-77, Gln-142 to His-148, Phe-232 to Gly-242, Pro-271 to Leu-278, Ser-340 to Asp-347, Pro-365 to Asp-371, Asp-398 to Leu-406, Arg-500 to Pro-505. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in uterine cancer and homology to prolylcarboxypeptidase indicates that the protein product of this gene would be useful
10 for diagnosis, treatment and prevention of diseases associated with the reproductive system including uterine cancer, as well as, cardiovascular diseases where prolylcarboxypeptidases primary substrate, angiotension, has its greatest effect. In addition, the putative location of prolylcarboxypeptidase within the lysosomal compartment of cells indicates that polynucleotides and polypeptides corresponding
15 to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylketonuria, galactosemia, porphyrias, and Hurler's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to
20 its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO:98 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
30 formula of a-b, where a is any integer between 1 to 1709 of SEQ ID NO:98, b is an integer of 15 to 1723, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:98, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 89

The translation product of this gene shares sequence homology with the human CGI-06 protein (See, e.g. Genbank Accession No.

- 5 gb|AAD27715.1|AF132940_1 (AF132940); all references available through this accession are hereby incorporated by reference herein). When tested against the myeloid cell line, U937, supernatants removed from cells containing this gene activated the GAS (gamma activation site) pathway. Thus, it is likely that this gene activates myeloid cells through the Jaks-STAT signal transduction pathway. The GAS
10 (gamma activation site) is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and
15 differentiation of cells.

The gene encoding the disclosed cDNA is believed to reside on chromosome 20. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 20.

- This gene is expressed primarily in various tumors including endometrial
20 tumors, adenocarcinoma, breast cancer, osteosarcoma, chondrosarcoma, uterine and pancreas tumors and to a lesser extent in embryonic tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, identification and treatment of many types of solid tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the major organs, expression of this gene at significantly higher or lower levels is routinely detected in
30 certain tissues or cell types (e.g., skeletal, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., breast milk, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such

a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 218 as residues: Pro-25 to Arg-31, Thr-52 to Val-63,
5 Asn-129 to Lys-135, Gln-197 to Trp-202, Thr-230 to Glu-236, Pro-242 to Tyr-248, Leu-280 to Pro-291, Ser-348 to Ser-356, Pro-362 to Gln-368, Thr-398 to His-406, Trp-430 to Leu-435, Glu-499 to Gly-504. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in solid tumors combined with the GAS-element
10 activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Representative uses are described in the "Hyperproliferative Disorders" and
15 "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain
20 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,
25 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

The protein is useful in modulating the immune response to aberrant
30 polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Additionally, the expression in hematopoietic cells and tissues indicates

that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:99 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2073 of SEQ ID NO:99, b is an integer of 15 to 2087, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:99, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 90

This gene is expressed primarily in brain medulloblastoma cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of brain medulloblastoma and other neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an

individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in medulloblastoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:100 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 737 of SEQ ID NO:100, b is an integer of 15 to 751, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:100, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 91

This gene maps to the chromosome X, and therefore, is used as a marker in linkage analysis for chromosome X.

Preferred polypeptides of the invention comprise the following amino acid sequence: CSVFPPSLWFYLPVFDGDVQ (SEQ ID NO: 489), GVSLPLLGDASQLGYLGV RDALEEALCLFSDVQLCAGRTSALFKAXRQGRSLQRILLPFVWLCPAPQRWSLQRQAGLLELRWAPPS SSFLAALFTPSSLGNGGRPSPSLTAXLQFDLRLLC (SEQ ID NO: 490), and/or VCRGFCC
 5 LLFGCALPPRGGVYRGRQASLNCGLHRVRVSWPLCLPPQASAMVGA PPPASLPXCSLISDCASNX (SEQ ID NO: 491). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in spleen from chronic lymphocytic leukemia patients.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia, and other immune disorders, particularly proliferative diseases. Similarly, polypeptides and antibodies directed to
 15 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,
 20 plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spleen from chronic lymphocytic leukemia patients
 25 indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders.

Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product in leukemia cells indicates a role in the
 30 regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or

other processes that may also suggest a usefulness in the treatment of cancer e.g., by boosting immune responses.

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, and leukemia. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:101 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1209 of SEQ ID NO:101, b is an integer of 15 to 1223, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:101, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 92

The translation product of this gene was shown to have homology to the human reverse transcriptase gene (See e.g., Genbank Accession No. gi|439877; all references available through this accession are hereby incorporated by reference herein).

Preferred polypeptides of the invention comprise the following amino acid sequence: MSHKHMRRSATSYYIRERQIKIIVRYHYTPIMTT (SEQ ID NO: 492), IRERQIKIIVRYHYTP (SEQ ID NO: 493), KKTCTMFIATLFT (SEQ ID NO: 494), SVASVFIP

LKVSVTKQFIFFXFFFFLRRSLAPAWVAERXTSQETKQNKKTPQLRGKVAHACDPITLGRRWEVGESL
EARSPTS (SEQ ID NO: 496) and/or EKIFAKHLSVKGL (SEQ ID NO: 495).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in microvascular endothelial cells.

- 5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, various diseases of the cardiovascular and circulatory systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
- 10 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., vascular, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid
- 15 and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution in microvascular endothelial cells combined with the homology to the conserved human gene for reverse transcriptase indicates that
- 20 polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders, particularly vascular disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Alternatively expression within microvascular tissue, a tissue
- 25 marked by proliferating cells, indicates that this protein may play a role in the regulation of cellular division. As such, this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early
- 30 hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue

differentiation and could again be useful in cancer therapy. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
 5 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:102 and may have been publicly available prior to conception
 10 of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 996 of SEQ ID NO:102, b is an
 15 integer of 15 to 1010, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:102, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 93

The translation product of this gene shares sequence homology with the
 20 Y43F4B.5 protein from *Caenorhabditis elegans* (See Genebank Accession No. gn|PID|e1247424 (AL021481)). Moreover, the translation product also shares homology to phosphoglucomutase proteins (See Genbank Accession No. emb|CAA16334.1| (AL021481)). Based on the sequence similarity, The translation product of this gene is expected to share at least some biological activities with
 25 phosphoglucomutase proteins. Such activities are known in the art, some of which are described elsewhere herein.

Preferred polypeptides of the invention comprise the following amino acid sequence: ARGKTVLFAFEEAIGYMCCPFVLDKDGVSAAVISAE LASFLATKNLSLSQQLKAIYVEYG
 YHITKASYFICHQETIKKLFENLRNYDGKNYPKACGKFEISAIRDLTTGYDDSQPDKKAVLPTSKSS
 30 QMITFTFANGGVATMRTSGTEPKIKYYAELCAPPGNSDPEQLKKELNELVSAIEEHFFQPQKYNLQPKAD (SEQ ID NO: 498), YMCCPFVLDKDGVSAAVISAE LASFLATKNLSLSQQLKAIYVEYGYHIT
 KASYFICHQETIKKLFENLRNYDGKNYPKACGKFEISAIRDLTTGYDDSQPDKKAVLPTSKSSQMIT
 FTFANGGVATMRTSGTEPKIKYYAELCAPPGNSDPEQLKKELNELVSAIEEHFFQPQKYNLQPKAD

(SEQ ID NO:497), DKDGVSAAVISAEASFL (SEQ ID NO: 499), RDLTTGYDDSQPD (SEQ ID NO: 500), KAVLPTSKSSQMITF (SEQ ID NO: 501), and/or TMRTSGTEPKIKYYAEL (SEQ ID NO: 502). Polynucleotides encoding these polypeptides are also provided.

5 The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

 This gene is expressed primarily in placenta, fetal spleen, and to a lesser extent in prostate, T-cells and neutrophils.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, various diseases of the immune and reproductive systems, including cancer. Similarly, polypeptides and antibodies directed to these polypeptides are
15 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and reproductive systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g.,
20 seminal fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred polypeptides of the present invention comprise immunogenic
25 epitopes shown in SEQ ID NO: 222 as residues: Leu-23 to Met-30. Polynucleotides encoding said polypeptides are also provided.

 The tissue distribution in fetal spleen indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the
30 "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or

activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the

5 natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to

10 transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other

15 blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and conditions, which include, but are

20 not limited to microvascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

25 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:103 and may have been publicly available prior to conception

30 of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1972 of SEQ ID NO:103, b is an integer of 15 to 1986, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:103, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 94

This gene is expressed primarily in activated monocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, various diseases and/or disorders of the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in activated monocytes indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency

diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:104 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1319 of SEQ ID NO:104, b is an integer of 15 to 1333, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:104, and where b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HWBBP10	209782 04/20/98	pCMVSPORT 3.0	11	899	1	899	66	66	130	1	26	27	56
1	HWBBP10	209782 04/20/98	pCMVSPORT 3.0	105	944	1	944	55	55	224	1	26	27	56
2	HWBDO80	209782 04/20/98	pCMVSPORT 3.0	12	1140	1	1140	166	166	131	1	22	23	41
3	HWHGU54	209782 04/20/98	pCMVSPORT 3.0	13	1445	1	1445	145	145	132	1	19	20	414
4	HYACI76	209782 04/20/98	pCMVSPORT 3.0	14	1208	1	1148	385	385	133	1	25	26	44
5	HBHMA23	209782 04/20/98	pSPORT1	15	1175	2	1175	71	71	134	1	24	25	197
5	HBHMA23	209782 04/20/98	pSPORT1	106	1172	1	1172	70	70	225	1	24	25	76
6	HCE3G20	209782 04/20/98	Uni-ZAP XR	16	2374	1	2350	57	57	135	1	42	43	45
7	HCEJP80	209782 04/20/98	Uni-ZAP XR	17	1595	1	1595	90	90	136	1	21	22	40
8	HCUDDD24	209782 04/20/98	ZAP Express	18	1287	89	1287	314	314	137	1	19	20	84

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
9	HDPTD15	209782 04/20/98	pCMVSPORT 3.0	19	1396	1	1396	223	223	138	1	18	19	200
10	HDPWU34	209782 04/20/98	pCMVSPORT 3.0	20	1277	860	1277	117	117	139	1	23	24	325
10	HDPWU34	209782 04/20/98	pCMVSPORT 3.0	107	427	1	427	111	111	226	1	16	17	44
11	HEOOV79	209782 04/20/98	pSPORT1	21	1781	1	1767	203	203	140	1	23	24	118
12	HFKET93	209782 04/20/98	Uni-ZAP XR	22	1491	1	1491	75	75	141	1	15	16	47
13	HFTDL56	209782 04/20/98	Uni-ZAP XR	23	1839	32	1838	93	93	142	1	22	23	519
14	HFJX44	209782 04/20/98	Lambda ZAP II	24	1384	1	1384	98	98	143	1	18	19	47
15	HKACU58	209782 04/20/98	pCMVSPORT 2.0	25	1681	1	1681	98	98	144	1	18	19	431
15	HKACU58	209782 04/20/98	pCMVSPORT 2.0	108	1708	69	1708	117	117	227	1	18	19	101
16	HKFBC53	209782 04/20/98	ZAP Express	26	1949	1	1906	41	41	145	1	18	19	442
16	HLDBQ19	209226 08/28/97	pCMVSPORT 3.0	109	1487	401	1487	534	534	228	1	22	23	132

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
16	HLDBQ19	97958 03/13/97	pCMVSPORT 3.0	110	1525	401	1480	534	534	229	1	22	23	66
17	HLTHR66	209782 04/20/98	Uni-ZAP XR	27	2286	1	2286	5	5	146	1	34	35	75
18	HLYBA69	209782 04/20/98	pSport1	28	530	1	530	89	89	147	1	29	30	51
19	HNTMX29	209782 04/20/98	pSport1	29	1296	756	1291	118	118	148	1	31	32	209
19	HNTMX29	209782 04/20/98	pSport1	111	552	1	552	18	18	230	1	18	19	72
20	HNTNC20	209782 04/20/98	pSport1	30	1979	1	1979	270	270	149	1	19	20	218
21	HNTNI01	209782 04/20/98	pSport1	31	1274	1	1114	306	306	150	1	33	34	49
22	HOHCK70	209782 04/20/98	pCMVSPORT 2.0	32	1531	1	1531	245	245	151	1	27	28	40
23	HSMBE69	209782 04/20/98	pSport1	33	2090	1	2090	69	69	152	1	18	19	107
24	HT4FW61	209782 04/20/98	Uni-ZAP XR	34	1006	31	1006	107	107	153	1	38	39	156
25	HYABK95	209782 04/20/98	pCMVSPORT 3.0	35	1787	1	1787	267	267	154	1	26	27	150

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
26	HYACE88	209782 04/20/98	pCMVSPORT 3.0	36	1201	1	1180	316	316	155	1	16	17	70
27	HOABR60	209782 04/20/98	Uni-ZAP XR	37	1896	1	903	45	45	156	1	16	17	490
27	HOABR60	209782 04/20/98	Uni-ZAP XR	112	925	1	903	45	45	231	1	16	17	293
28	HAGCT73	209782 04/20/98	Uni-ZAP XR	38	1152	1	1152	119	119	157	1	30	31	31
29	HAPOM45	209782 04/20/98	Uni-ZAP XR	39	1017	34	1017	98	98	158	1	31	32	115
30	HCEJQ69	209782 04/20/98	Uni-ZAP XR	40	1777	1	1777	39	39	159	1	26	27	380
31	HAGFI62	209782 04/20/98	Uni-ZAP XR	41	1003	368	992	429	429	160	1	28	29	91
32	HAGGS43	209782 04/20/98	Uni-ZAP XR	42	1201	1	1201	62	62	161	1	25	26	44
33	HBJHP03	209852 05/07/98	Uni-ZAP XR	43	1176	1	1176	185	185	162	1	20	21	45
34	HCHPF68	209852 05/07/98	pSport1	44	569	1	569	186	186	163	1	36	37	128
35	HDPJF37	209852 05/07/98	pCMVSPORT 3.0	45	986	1	986	196	196	164	1	23	24	57

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
36	HSDEZ20	209852 05/07/98	Uni-ZAP XR	46	1540	1	1540	66	66	165	1	41	42	97
37	HTEKU58	209852 05/07/98	Uni-ZAP XR	47	792	73	792	93	93	166	1	30	31	59
38	HLTBL58	209852 05/07/98	Uni-ZAP XR	48	1497	1	1497	26	26	167	1	20	21	42
39	HPWDJ42	209852 05/07/98	Uni-ZAP XR	49	1340	1	1340	149	149	168	1	18	19	54
39	HPWDJ42	209852 05/07/98	Uni-ZAP XR	113	1340	1	1340	149	149	232	1	21	22	54
39	HPWDJ42	209852 05/07/98	Uni-ZAP XR	114	813	1	813	161	161	233	1	18	19	47
40	HRACD15	209852 05/07/98	pCMVSPORT 3.0	50	1539	24	1539	252	252	169	1	40	41	53
40	HRACD15	209852 05/07/98	pCMVSPORT 3.0	115	1681	24	1453	252	252	234	1	40	41	53
41	HSIAC80	209852 05/07/98	Uni-ZAP XR	51	1423	1	1423	178	178	170	1	17	18	53
42	HAGFD18	209852 05/07/98	Uni-ZAP XR	52	1364	94	1364	261	261	171	1	21	22	48
43	HMTAT59	209852 05/07/98	pCMVSPORT 3.0	53	2288	501	2276	301	301	172	1	14	15	224

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
44	HDTGC86	209852 05/07/98	pCMVSPORT 2.0	54	1512	1	1512	351	351	173	1	27	28	200
45	HAGDI35	209852 05/07/98	Uni-ZAP XR	55	1357	1	1338	318	318	174	1	25	26	93
46	HELHN47	209852 05/07/98	Uni-ZAP XR	56	1989	883	1989	778	778	175	1	30	31	404
47	HPRBC80	209852 05/07/98	Uni-ZAP XR	57	2543	1245	2543	94	94	176	1	30	31	387
47	HPRBC80	209852 05/07/98	Uni-ZAP XR	116	2052	275	2032	404	404	235	1	26	27	69
48	HAQAR23	209852 05/07/98	Uni-ZAP XR	58	777	66	777	92	92	177	1	19	20	145
49	HAIFL18	209852 05/07/98	Uni-ZAP XR	59	879	1	879	274	274	178	1	29	30	140
50	HJPAY76	209852 05/07/98	Uni-ZAP XR	60	1161	1	1161	134	134	179	1	21	22	127
51	HUSXE77	209852 05/07/98	pSport1	61	687	1	687	156	156	180	1	20	21	146
52	HUFEF62	209852 05/07/98	pSport1	62	518	1	518	190	190	181	1	28	29	68
52	HUFEF62	209852 05/07/98	pSport1	117	539	1	539	182	182	236	1	28	29	68

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
53	HTWJK32	209852 05/07/98	Lambda ZAP II	63	911	211	911	376	376	182	1	20	21	51
54	HTWDF76	209852 05/07/98	pSport1	64	963	1	963	316	316	183	1	24	25	85
55	HTPBN68	209852 05/07/98	Uni-ZAP XR	65	1001	1	1001	429	429	184	1	20	21	191
56	HTOIY21	209852 05/07/98	Uni-ZAP XR	66	1558	1	1558	91	91	185	1	14	15	231
57	HTLDD53	209852 05/07/98	Uni-ZAP XR	67	1322	1	1322	162	162	186	1	25	26	68
58	HTLFG05	209852 05/07/98	Uni-ZAP XR	68	865	1	717	137	137	187	1	30	31	211
58	HTLFG05	209852 05/07/98	Uni-ZAP XR	118	882	1	882	137	137	237	1	30	31	67
59	HDPXR23	209852 05/07/98	pCMVSPORT 3.0	69	1150	20	1150	49	49	188	1	20	21	90
59	HDPXR23	209852 05/07/98	pCMVSPORT 3.0	119	1193	1	1189	95	95	238	1	20	21	90
60	HSIAC45	209852 05/07/98	Uni-ZAP XR	70	1398	1	1398	12	12	189	1	23	24	62
61	HSRGW16	209853 05/07/98	Uni-ZAP XR	71	1557	180	1007	72	72	190	1	12	13	295

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
61	HSRGW16	209853 05/07/98	Uni-ZAP XR	120	1338	1	1338	170	170	239	1	47	48	140
62	HSSJC35	209853 05/07/98	Uni-ZAP XR	72	1163	1	1163	55	55	191	1	30	31	295
62	HSSJC35	209853 05/07/98	Uni-ZAP XR	121	1183	1	1183	66	66	240	1	30	31	37
63	HTEAX23	209853 05/07/98	Uni-ZAP XR	73	1486	1	1486	72	72	192	1	20	21	338
64	HTGCH22	209853 05/07/98	Uni-ZAP XR	74	1553	1	1553	12	12	193	1	29	30	78
65	HTJMA95	209853 05/07/98	pCMVSPORT 2.0	75	1650	198	1569	527	527	194	1	22	23	181
66	HHEAA08	209853 05/07/98	pCMVSPORT 3.0	76	2150	1	2150	88	88	195	1	38	39	79
66	HHEAA08	209853 05/07/98	pCMVSPORT 3.0	122	615	1	615		311	241	1			20
67	HBQAA49	209853 05/07/98	Lambda ZAP II	77	1592	1	1592	197	197	196	1	37	38	69
68	HDPBI32	209853 05/07/98	pCMVSPORT 3.0	78	1579	598	1184	103	103	197	1	30	31	271
68	HDPBI32	209853 05/07/98	pCMVSPORT 3.0	123	587	1	587	51	51	242	1	35	36	138

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
69	HBIBF16	209853 05/07/98	Uni-ZAP XR	79	1396	1	1396	15	15	198	1	35	36	51
70	HBCAY05	209853 05/07/98	Uni-ZAP XR	80	1230	576	1209	627	627	199	1	22	23	71
71	HCUCK44	209853 05/07/98	ZAP Express	81	1139	573	1133	593	593	200	1	30	31	60
72	HCE2W56	209853 05/07/98	Uni-ZAP XR	82	1409	1	1409	61	61	201	1	21	22	143
73	HCWAG01	209853 05/07/98	ZAP Express	83	714	1	714	192	192	202	1	25	26	148
74	HLDBY02	209853 05/07/98	pCMVSPORT 3.0	84	1097	1	1097	326	326	203	1	30	31	36
75	HDRMI82	209853 05/07/98	pSport1	85	1931	540	1900	170	170	204	1	25	26	406
75	HDRMI82	209853 05/07/98	pSport1	124	1379	1	1357	328	328	243	1	30	31	175
76	HEPCU48	209853 05/07/98	Uni-ZAP XR	86	1092	1	1092	98	98	205	1	26	27	91
77	HDPRK33	209853 05/07/98	pCMVSPORT 3.0	87	578	1	573	99	99	206	1	44	45	101
78	HKGAX42	209853 05/07/98	pSport1	88	699	1	699	69	69	207	1	18	19	50

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
79	HLMMAZ95	209853 05/07/98	Uni-ZAP XR	89	1126	7	1126	187	187	208	1	33	34	161
80	HLMFC07	209853 05/07/98	Lambda ZAP II	90	1037	1	1037	203	203	209	1	17	18	227
80	HLMFC07	209853 05/07/98	Lambda ZAP II	125	1268	1	1268	203	203	244	1	30	31	39
81	HL2AG87	209853 05/07/98	Uni-ZAP XR	91	1316	1	1316	110	110	210	1	37	38	351
82	HKGCO27	209853 05/07/98	pSport1	92	1021	1	1021	313	313	211	1	26	27	93
82	HKGCO27	209853 05/07/98	pSport1	126	1311	1	1311	57	57	245	1	26	27	47
83	HLDCE79	209853 05/07/98	pCMVSPORT 3.0	93	1260	1	1260	342	342	212	1	63	64	101
83	HLDCE79	209853 05/07/98	pCMVSPORT 3.0	127	1249	1	1249	298	298	246	1	30	31	34
84	HERAD40	209853 05/07/98	Uni-ZAP XR	94	990	1	990	85	85	213	1	38	39	98
85	HFOX55	209853 05/07/98	pSport1	95	1710	1	1710	138	138	214	1	34	35	81
86	HFBVGZ42	209853 05/07/98	pBluescript	96	781	1	781	71	71	215	1	22	23	188

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
87	HNHAF39	209853 05/07/98	Uni-ZAP XR	97	1113	1	1113	332	332	216	1	30	31	44
88	HNTSW57	209853 05/07/98	pSport1	98	1723	181	1723	19	19	217	1	21	22	515
88	HNTSW57	209853 05/07/98	pSport1	128	1660	1	1660	38	38	247	1	21	22	490
89	HOGCK20	209853 05/07/98	pCMVSPORT 2.0	99	2087	1	2087	57	57	218	1	23	24	522
89	HOGCK20	209853 05/07/98	pCMVSPORT 2.0	129	2075	1	2054		53	248	1	22	23	554
90	HMDAL49	209853 05/07/98	Uni-ZAP XR	100	751	1	751	52	52	219	1	22	23	52
91	HLYES38	209853 05/07/98	pSport1	101	1223	1	1223	69	69	220	1	22	23	73
92	HMECK83	209853 05/07/98	Lambda ZAP II	102	1010	1	1010	50	50	221	1	28	29	54
93	HSHAX21	209853 05/07/98	Uni-ZAP XR	103	1986	1	1986	177	177	222	1	13	14	72
94	HMQAG66	209853 05/07/98	Uni-ZAP XR	104	1333	1	1333	657	657	223	1	24	25	69

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization

probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID
5 NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or
10 deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide
15 sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily
20 be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a
25 suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed
30 herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

5 The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

10 The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during
15 recombinant production.

 The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40
20 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

25 Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the
30 information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of

these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95%

"identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence

that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid.

These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

- 5 As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity.
- 10 Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.
- 15 If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from
- 20 the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and
- 25
- 30

C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

5 For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

15 The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

25 Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an

organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level.

Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

5 Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988
10 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

15 Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over
20 the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in
25 activity from wild-type.

 Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted
30 form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic

activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, 5 inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

10 The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these 15 positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. 20 For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are 25 surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions 30 involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of

the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention
5 include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for
10 example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of
15 charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev.
20 Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid
25 substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1
30 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is

1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

Polynucleotide and Polypeptide Fragments

5 In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt
10 in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

15 Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-
20 1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini.
25 Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

 In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the
30 deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the

invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these

fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

5 **Fusion Proteins**

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second
10 protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous
15 functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the
20 polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-
25 polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the
30

IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion
5 proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified,
10 would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol.
15 Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue,
20 Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767
25 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

30 The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral

vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

5 The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

10 The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding
15 portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin
20 resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as
25 CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-
9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A,
pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and
30 ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1

and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids

containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or

translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more
5 restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

10 The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a
15 unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological
20 samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with
25 one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a
30 particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the

present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as
5 molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an
10 immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

15 A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene
20 expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I , ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), and fluorescent labels, such as fluorescein and rhodamine, and
25 biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which
30 emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as

deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I , ^{112}In , $^{99\text{m}}\text{Tc}$), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of $^{99\text{m}}\text{Tc}$. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic

anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, 5 Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide 10 or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ 15 rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of 20 T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic 25 and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or 30 IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response.

Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae,

Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable

of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell
5 membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

10 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations,
15 polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

20 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers.
25 The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

30 Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a

candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95%

identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under
5 stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which
10 comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide
15 sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of
20 at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

25 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule
30 comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X
5 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and
10 determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence
15 selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

20 A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X
25 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological
30 sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least

one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of

positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence
5 at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

10 Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino
15 acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted
20 protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the
25 amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in
30 the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA

clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of

the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1
5 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an
10 isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

15

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector.
20 Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being
25 isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

30

<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited</u>
<u>Plasmid</u>	
Lambda Zap	pBluescript (pBS)
Uni-Zap XR	pBluescript (pBS)
Zap Express	pBK
lafmid BA	plafmid BA

pSport1	pSport1
pCMVSPORT 2.0	pCMVSPORT 2.0
pCMVSPORT 3.0	pCMVSPORT 3.0
pCR [®] 2.1	pCR [®] 2.1

- 5 Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128,256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Altting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Altting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are
- 10 commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3
- 15 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.
- 20 Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University,
- 25 NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al.,
- 30 Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ^{32}P - γ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 μl of reaction

mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific

to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

5 **Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide**

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

10 **Example 3: Tissue Distribution of Polypeptide**

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling
15 system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

20 Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed
25 according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This
30 primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and

hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1

mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The

origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without
5 mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive
10 Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4
15 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using
20 a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the
25 above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus

Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector
5 contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak
10 *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such
15 as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

20 Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP)
25 to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a
30 commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("GeneClean" BIO 101 Inc., La Jolla, Ca.).

5 The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by
10 gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method
15 described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the
20 transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

25 After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell
30 culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then

resuspended in a microcentrifuge tube containing 200 μ l of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

5 To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and
10 cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μ Ci of 35 S-methionine and 5 μ Ci 35 S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

15 Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

 The polypeptide of the present invention can be expressed in a mammalian
20 cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient
25 transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

 Suitable expression vectors for use in practicing the present invention include,
30 for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109),

pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

5 Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

 The transfected gene can also be amplified to express large amounts of the
10 encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker
15 is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the
20 production of proteins.

 Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the
25 CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

30 Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five μ g of the expression plasmid pC6 is cotransfected with 0.5 μ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see
5 also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create
10 chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

15 Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion
20 can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will
25 not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

30 Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGC
CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAAA
CCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTACATGCGTGGT
GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG
5 ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA
ACCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG
10 GTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGT
GGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCT
CCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCA
TGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
15 GTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a
20 polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

25 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell
30 Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in

any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

5 The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as
10 described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

 Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method
15 makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody
20 whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

 It will be appreciated that Fab and F(ab')₂ and other fragments of the
25 antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic
30 chemistry.

 For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced

using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening

10 Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution
15 (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered
20 Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x
25 Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods
30 described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45

minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following
 5 tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells
 10 first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl₂ (anhyd); 0.00130 mg/L CuSO₄·5H₂O; 0.050 mg/L of Fe(NO₃)₃·9H₂O; 0.417 mg/L of FeSO₄·7H₂O;
 15 311.80 mg/L of KCl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄·H₂O; 71.02 mg/L of Na₂HPO₄; .4320 mg/L of ZnSO₄·7H₂O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of
 20 Oleic Acid; 0.010 mg/L of Palmitic Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0
 25 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H₂O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of
 30 Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319

mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of

5 Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for

10 endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

15 On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the

20 polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

25 **Example 12: Construction of GAS Reporter Construct**

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The

30 binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in
5 many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus
10 upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table
15 below. (Adapted from review by Schindler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved
20 cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

25 Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway
30 can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u> <u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	<u>STATS</u>	<u>GAS(elements) or ISRE</u>
	<u>IFN family</u>						
5	IFN-a/B	+	+	-	-	1,2,3	ISRE
	IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
	Il-10	+	?	?	-	1,3	
	<u>gp130 family</u>						
10	IL-6 (Pleiotrophic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
	Il-11(Pleiotrophic)	?	+	?	?	1,3	
	OnM(Pleiotrophic)	?	+	+	?	1,3	
	LIF(Pleiotrophic)	?	+	+	?	1,3	
	CNTF(Pleiotrophic)	-/+	+	+	?	1,3	
15	G-CSF(Pleiotrophic)	?	+	?	?	1,3	
	IL-12(Pleiotrophic)	+	-	+	+	1,3	
	<u>g-C family</u>						
	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
20	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
	IL-13 (lymphocyte)	-	+	?	?	6	GAS
	IL-15	?	+	?	+	5	GAS
25	<u>gp140 family</u>						
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5	GAS
	GM-CSF (myeloid)	-	-	+	-	5	GAS
30	<u>Growth hormone family</u>						
	GH	?	-	+	-	5	
	PRL	?	+/-	+	-	1,3,5	
	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
35	<u>Receptor Tyrosine Kinases</u>						
	EGF	?	+	+	-	1,3	GAS (IRF1)
	PDGF	?	+	+	-	1,3	
	CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

10 5':GCGCCTCGAGATTTCCTCCCGAAATCTAGATTTCCTCCCGAAATGATTTCCTCCCGAAATGATTTCCTCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTGTCAAAGCCTAGGC:3' (SEQ ID NO:4)

15 PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

20 5':CTCGAGATTTCCTCCCGAAATCTAGATTTCCTCCCGAAATGATTTCCTCCCGAAATGATTTCCTCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCC
 25 AGGCTTTTGTCAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not
5 contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance
10 gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS
15 with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-
20 2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

25 The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells
30 (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

The above protocol may be used in the generation of both transient, as well as, stable transfected cells, which would be apparent to those of skill in the art.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^7 U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing

10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM
5 KCl, 375 uM Na₂HPO₄·7H₂O, 1 mM MgCl₂, and 675 uM CaCl₂. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400
10 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in
15 the 96-well plate (or 1×10^5 cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant
20 according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes,
25 EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or
30 differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor).

The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

- 5 The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

- 10 Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

- 15 To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

- 20 PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

- 25 Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

- 30 To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

(Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count
5 the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

Add 200 μ l of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 μ l supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR
10 can be used, such as 50 ng/ μ l of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

15 NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of
20 apoptosis (NF- κ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target
25 genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in
30 treating diseases. For example, inhibitors of NF- κ B could be used to treat those

diseases related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCC
ATCCCCGCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGA
CTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTA
TTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA
GCTT:3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly,

the NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with Sall and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly,
 5 the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

10 As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x
 15 dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room
 20 temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

25 Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25

16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small

molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star
5 black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To
load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate
10 is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are
re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml
fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension.
15 The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed
twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100
ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed
once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final
volume.

20 For a non-cell based assay, each well contains a fluorescent molecule, such as
fluo-4. The supernatant is added to the well, and a change in fluorescence is
detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the
following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4
25 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm;
and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an
extracellular signaling event which has resulted in an increase in the intracellular
Ca⁺⁺ concentration.

30 **Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase
Activity**

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies.

- 5 In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation
10 of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

- 15 Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

- 20 Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 µl of cell culture grade type I collagen (50 µg/ml), gelatin (2%) or
25 polylysine (50 µg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar
30 Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen

Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium.

- 5 Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇ and a cocktail of protease inhibitors (# 1836170) obtained from
- 10 Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on
- 15 ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described

20 here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

25 PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂⁺ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride,

30 pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the

components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

- 5 Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-
10 POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

- Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound
15 peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

- 20 As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other
25 molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

- Specifically, assay plates are made by coating the wells of a 96-well ELISA
30 plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G

plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1

and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place

of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale).

Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the

scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1
5 $\mu\text{g/kg/day}$ to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day , and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 $\mu\text{g/kg/hour}$, either by 1-
10 4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention
15 are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to
20 modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or
25 microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or
30 poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al.,

Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar

alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an

individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

5 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the
10 presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

 The amphotropic pA317 or GP+am12 packaging cells are grown in tissue
15 culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

20 Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the
25 media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is
30 produced.

 The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 27: Method of Treatment Using Gene Therapy - In Vivo

5 Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter
10 or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord.
15 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The
20 polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or
25 precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

30 The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in

the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an

aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

5 The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

10 Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the
15 knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A
20 time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper
25 dosages and other treatment parameters in humans and other animals using naked DNA.

Example 28: Transgenic Animals.

The polypeptides of the invention can also be expressed in transgenic animals.
30 Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a

specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (i.e.,
5 polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus
10 mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science
15 259:1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

20 Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

25 The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and
30 activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the

particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 10 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening 15 may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also 20 be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding 25 strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given 30 integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the

transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Example 29: Knock-Out Animals.

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (*E.g.*, see Smithies et al., *Nature* 317:230-234 (1985); Thomas & Capecchi, *Cell* 51:503-512 (1987); Thompson et al., *Cell* 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (*e.g.*, see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (*e.g.*, knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (*i.e.*,

animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

5 The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or

10 intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and

15 Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For

20 example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological

25 function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>198</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit April 20, 1998	Accession Number 209782
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: right; margin-bottom: 5px;">For receiving Office use only</div> <div style="margin-bottom: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div>	<div style="text-align: right; margin-bottom: 5px;">For International Bureau use only</div> <div style="margin-bottom: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div>
--	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>199</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit August 28, 1997	Accession Number 209226
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For receiving Office use only</div> <div style="margin-bottom: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div>	<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For International Bureau use only</div> <div style="margin-bottom: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div>
---	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>200</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>March 13, 1997</u>	Accession Number <u>97958</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For receiving Office use only</div> <div style="margin-bottom: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div>	<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For International Bureau use only</div> <div style="margin-bottom: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; height: 40px; margin-top: 5px;"></div>
---	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>201</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 7, 1998</u>	Accession Number <u>209852</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer</div>	<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer</div>

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>204</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 7, 1998</u>	Accession Number <u>209853</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer</div>	<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer</div>

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

- 5 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both
- 10 incorporated herein by reference in their entireties.

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group
5 consisting of:
 - (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a
10 polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - 15 (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X,
20 having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
 - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
 - 25 (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.

5 3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

10 4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

15 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

20 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

25 7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

30 9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - 5 (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
 - (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - 10 (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - 15 (g) a variant of SEQ ID NO:Y;
 - (h) an allelic variant of SEQ ID NO:Y; or
 - (i) a species homologue of the SEQ ID NO:Y.
12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.
- 20
13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.
- 25
14. A recombinant host cell that expresses the isolated polypeptide of claim 11.
15. A method of making an isolated polypeptide comprising:
- (a) culturing the recombinant host cell of claim 14 under conditions such that
 - 30 said polypeptide is expressed; and
 - (b) recovering said polypeptide.

16. The polypeptide produced by claim 15.
17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount
5 of the polypeptide of claim 11 or the polynucleotide of claim 1.
18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or absence of a mutation in the polynucleotide of
10 claim 1; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.
19. A method of diagnosing a pathological condition or a susceptibility to
15 a pathological condition in a subject comprising:
- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.
20
20. A method for identifying a binding partner to the polypeptide of claim
11 comprising:
- (a) contacting the polypeptide of claim 11 with a binding partner; and
- (b) determining whether the binding partner effects an activity of the
25 polypeptide.
21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.
22. A method of identifying an activity in a biological assay, wherein the
30 method comprises:
- (a) expressing SEQ ID NO:X in a cell;
- (b) isolating the supernatant;

- (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

1

<110> Human Genome Sciences, Inc. et al.

<120> 94 Human secreted proteins

<130> PZ029PCT

<140> Unassigned

<141> 1999-06-13

<150> 60/089,508

<151> 1998-06-16

<150> 60/089,507

<151> 1998-06-16

<150> 60/089,510

<151> 1998-06-16

<150> 60/089,509

<151> 1998-06-16

<150> 60/090,112

<151> 1998-06-22

<150> 60/090,113

<151> 1998-06-22

<160> 502

<170> PatentIn Ver. 2.0

<210> 1

<211> 733

<212> DNA

<213> Homo sapiens

<400> 1

gggatccgga gcccaaactct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg	60
aattcgaggg tgcaccgtca gtcttctctt tcccccaaaa acccaaggac accctcatga	120
tctcccgac tcttgaggtc acatgcgtgg tgggtggacgt aagccacgaa gaccctgagg	180
tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg	240
aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact	300
ggctgaatgg caaggagtac aagtgaagg tctccaacaa agccctccca acccccatcg	360
agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc	420
catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct	480
atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagAAC aactacaaga	540
ccacgcctcc cgtgctggac tccgacggct ccttcttctt ctacagcaag ctcaccgtgg	600
acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc	660
acaaccacta cagcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc	720
gactctagag gat	733

<210> 2

<211> 5

<212> PRT

<213> Homo sapiens

<220>

<221> Site

<222> (3)

<223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2
 Trp Ser Xaa Trp Ser
 1 5
 <210> 3
 <211> 86
 <212> DNA
 <213> Homo sapiens

<400> 3
 gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc 60
 cccgaaatat ctgccatctc aattag 86

<210> 4
 <211> 27
 <212> DNA
 <213> Homo sapiens

<400> 4
 gcggcaagct ttttgcaaag cctagggc 27

<210> 5
 <211> 271
 <212> DNA
 <213> Homo sapiens

<400> 5
 ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60
 aaatatctgc catctcaatt agtcagcaac catagtcccc cccctaactc cgcccatccc 120
 gccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa ttttttttat 180
 ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240
 ttttgagggc ctaggctttt gcaaaaagct t 271

<210> 6
 <211> 32
 <212> DNA
 <213> Homo sapiens

<400> 6
 gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7
 <211> 31
 <212> DNA
 <213> Homo sapiens

<400> 7
 gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8
 <211> 12
 <212> DNA
 <213> Homo sapiens

<400> 8
 ggggactttc cc 12

<210> 9
 <211> 73

<212> DNA

<213> Homo sapiens

<400> 9

gcggcctcga	ggggactttc	ccggggactt	tccggggact	ttccgggact	ttccatcctg	60
ccatctcaat	tag					73

<210> 10

<211> 256

<212> DNA

<213> Homo sapiens

<400> 10

ctcgagggga	ctttcccg	gactttcc	ggactttcc	ggactttcca	tctgccatct	60
caattagtca	gcaaccatag	tcccgccct	aactccgccc	atcccgcgcc	taactccgcc	120
cagttccgcc	cattctccgc	cccatggctg	actaattttt	tttatttatg	cagaggccga	180
ggccgcctcg	gcctctgagc	tattccagaa	gtagtgagga	ggcttttttg	gaggcctagg	240
cttttgcaaa	aagctt					256

<210> 11

<211> 899

<212> DNA

<213> Homo sapiens

<400> 11

ccacgcgtcc	ggaaaaagta	caagcccctc	tcaaatgggt	caagtttcaa	atattagacc	60
cacccatggc	aaagacagat	tttagtataa	tactcctaaa	actacactgt	cttttttttt	120
tttctgtcat	aagtgtgcat	tgtgctcagt	catttatttc	agtgacccaa	acagagccca	180
gtccagctgt	ttgtattttc	cctgcagtgg	gaagtggact	agggccatgt	gactaagaaa	240
gccagcctgg	gggtgtctt	ttcacctaca	gatgttttaa	tgtgcttaac	attatccaat	300
actagaacac	gagatagtct	aaataccaca	gcaggatctg	attagctttt	tcagatcact	360
gcctttatatt	gctgtttgca	aaaaagctta	atccagtgtc	agagatcagg	cttcctgctg	420
agccctgggg	tagtttctct	cattctttgt	gttcacagt	gcaggcgta	gtgagcagat	480
tcctctctct	cctaaattaa	agctgtaaag	tagtaactgt	agtagcaagg	gataaaagaga	540
aggaagaaaa	cccaagggaa	aaaagaagac	tgtctattca	taccaagtag	tttccttgat	600
atacacaaaa	gaaagagttt	ctaatatgaa	ttcataaata	ctgacctcag	tgtctcttct	660
actcagtgc	cagctattaa	gttttattag	gtttcagttg	taactacttt	gtgtggatat	720
atgttacgtt	tttcatattt	atcctactca	atcaatctca	gttttaccag	aagaattaca	780
tttattagcc	ataacagtgg	cccttctctt	attcttttca	gggctgatat	ctttttttatt	840
catgagattt	caaaaagaac	tatcaccacc	actaacaaaa	aaaaaaaaaa	aaaaaaaaaa	899

<210> 12

<211> 1140

<212> DNA

<213> Homo sapiens

<400> 12

cccacgcgtc	cgctgatgtt	attagcagca	taaggcagtc	atcgatgagg	cttgaggggg	60
ccttctgttg	gggtcacgg	cacctttcca	cagtacagga	cttcgaactt	ctgagagttg	120
taaaaggcgt	cctcattatc	tttgctgggt	ttggcatcct	ctttcatggc	cgmttttagat	180
aattgcctta	tgtgtcta	aacatcagga	acctggctgg	ggtctgtggc	gcggaaaacg	240
tggcaggcca	tctgcgactc	ggggctgc	ggctgcgcct	tgatcaggt	ggcaaatag	300
gtgaggtcgt	ggctgtttg	gatgaagcgc	gagatatgct	gcgccttg	ctcgaagatg	360
aataccgcgc	ggttgggctg	cgtggccgac	ggactagtgc	ccccgagggc	cccagcgccc	420
ggcgcgggga	cgcaacgcag	gaagggcgcg	ctgagcacca	ggatcacctc	tcgggcccgc	480
ggcgcggcgc	agccgcccgc	ctcgggcttc	tggtgcgc	tgcggtatct	ggccatgagc	540
cagggcagca	taggcagcgt	ggtcctgtgg	tccaggcacg	accccccaac	gtaccacagc	600
cggaaaccgt	tatcgcttgg	cttcccgggg	ccgggctgag	ctgagacgcc	cggctcgggc	660
tccaggggggt	gcgggaacgg	ctcatcctga	atgcagctgg	gcggctycat	aactctcgcc	720

4

tcaccagggc	accgcgagg	ccggccgggc	gcaccgcgc	ccccactccc	gcgcagaagg	780
cgccgcccga	actgtgccaa	ctgccgcacc	gggctyccgc	gcctgacctg	gagcggcgcg	840
accccgaact	ccgcgcttca	gcagccctgc	cccatgcagc	acttccacgg	gcgcggctcg	900
gaggctccgg	cgccgggcac	cgaggcaagc	gcccggcagg	cgagggcggg	ttaaatgggc	960
atcctcctcc	tcgggctggc	gcctcgggca	ggacctcccc	ttcctccgtc	gcgggtttgc	1020
agggtcagag	gaccacgccc	aggggtcccc	cgccgcgtct	agaggatccc	tcgaggggcc	1080
caagcttasg	cgtgcatgsg	acgtcatagc	taatctccct	ataggagatt	gcaaaagggg	1140

<210> 13

<211> 1445

<212> DNA

<213> Homo sapiens

<400> 13

ggaaggctgc	aggaccagga	ccgaaaaagg	actaggaggc	tgggatcagc	aacaactggg	60
gaaggccaag	gaagactgac	ctgaggggaa	aggagaaact	ggggagggtg	ggtctactac	120
tcaacaggat	attcttcaag	gaaaatgaac	cccacactag	gcctggccat	ttttctggct	180
gttctcctca	cggtgaaagg	tcttctaaag	ccgagcttct	caccaaggaa	ttataaagct	240
ttgagcgagg	tccaaggatg	gaagcaaagg	atggcagcca	aggagcttgc	aaggcagaac	300
atggacttag	gctttaagct	gctcaagaag	ctggcctttt	acaaccctgg	caggaacatc	360
ttcctatccc	ccttgagcat	ctctacagct	ttctccatgc	tgtgcctggg	tgcccaggac	420
agcaccctgg	acgagatcaa	gcaggggttc	aacttcagaa	agatgccaga	aaaagatcct	480
catgagggct	tccattacat	catccacgag	ctgaccacga	agaccagga	cctcaaactg	540
agcattggga	acacgctgtt	cattgaccag	aggctgcagc	cacagcgtaa	gtttttggaa	600
gatgccaaag	acttttacag	tgccgaaacc	atccttacca	actttcagaa	tttggaatg	660
gctcagaagc	agatcaatga	ctttatcagt	caaaaaaccc	atgggaaaat	taacaacctg	720
atcgagaata	tagaccccg	cactgtgatg	cttcttgcaa	attatatatt	ctttcgagcc	780
aggtggaaac	atgagtttga	tccaaatgta	actaaaggag	aagatttctt	tctggagaaa	840
aacagttcag	tcaagggtgc	catgatgttc	cgtagtgcca	tataccaagt	tggctatgac	900
gataagctct	cttgacccat	cctggaaaata	ccctaccaga	aaaatatcac	agccatcttc	960
atccttccct	atgagggcaa	gctgaagcac	ttggagaagg	gattgcaggt	ggacactttc	1020
tccagatgga	aaacattact	gtcacgcagg	gtcgtagacg	tgtctgtacc	cagactccac	1080
atgacgggca	ccttcgacct	gaagaagact	ctctcctaca	taggtgtctc	caaaatcttt	1140
gaggaacatg	gtgatctcac	caagatcgcc	cctcatcgca	gcctgaaagt	ggcgagggct	1200
gtgcacaagg	ctgagctgaa	gatggatgag	aggggtacgg	aagggggcgc	tggcaccgga	1260
gcacagactc	tgcccatgga	gacaccactc	gtcgtcaaga	tagacaaacc	ctatctgctg	1320
ctgattttaca	gcgagaaaaat	accttccgtg	ctcttcctgg	gaaagattgt	taaccctatt	1380
ggaaaataaa	ggagaattcc	tgcttgccac	agaccccgaa	aaaaaaaaaa	aaaaagggcg	1440
gccgc						1445

<210> 14

<211> 1208

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (9)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (59)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (79)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (814)

<223> n equals a,t,g, or c

<400> 14

tagcgcggn	gatccattcc	ccagaacact	ataccctagc	tttcaaaact	attagtgcnt	60
ataaaggtcg	cctcaggtnc	ggtcgaattc	ccggctcgacc	cacgcgtccg	ctagaaagag	120
aggtagtgct	ctgcagggcc	acgggaggac	tcagtgcgca	cttgaaagca	tcaaacacag	180
tggaggggctc	atacgggggtg	ctcagtagat	gggcgcacat	ttttatagaa	tactgaggcc	240
cagagagggga	agggtgtcttg	tctgtggctg	catgggggct	cagtgggaaa	gccgggacta	300
aaagctggcc	ccaggctagc	tttgtgccag	gccatcctgc	tcttacacag	gggctgagaa	360
ccaggggcag	cccaggagtc	ctggatgggg	cagcagtcac	gttgatggg	gctgggggtg	420
tggctctccc	tttctgggct	ctcaggagtg	gtcagggcta	gccccagatc	tcccagacca	480
agaagaggag	cagcctgttg	ggagacactc	atgccctgac	atgagtcagt	gcatcaagag	540
aggccatcag	ccagtgggat	tcagcaagca	tgccctggcg	tgccctggtag	gggtgtgccc	600
atgggaggaa	gagaagagga	gctgccaccc	atttggggcc	ytccctctct	gggtccctcag	660
atttgccctt	cagcccarag	tctatgaaga	ccccgcggcc	ctggatgggtg	gggaggaggg	720
catggacatc	wttaccacaca	ttctggcctt	ggcaccccg	ctcctgaaag	actctggtag	780
tatcttctta	gaagtggacc	caaggcaccc	ggancttgct	agcagctggc	ttcagagccg	840
gcctgacctg	taccttaatc	ttgtggctgt	gcgcagggac	ttctgtggga	ggccccgggt	900
cctgcataatc	cggaggtctg	ggccatagca	tggtctgcc	gtggatgcct	tgctcagtgcc	960
gccagcctga	ccagagggga	ggtggatggc	actttccaga	gcccagggtc	ttatggcatt	1020
tcccaggggt	ctgtgatttc	cccatgctct	gcatttctag	gatatttcta	ggacacctgg	1080
attggctcca	tcacatcaga	gtggctgagg	gcagttgctc	tgtgttggtg	aaattgctgt	1140
gggggtatcg	ggggatatgg	ccagtaaaagt	attgagagac	taacaaaaaa	aaaaaaaaaa	1200
aaactcga						1208

<210> 15

<211> 1175

<212> DNA

<213> Homo. sapiens

<400> 15

gagcgggccg	aggactccag	cgtgcccagg	tctggcatcc	tgcaacttgc	gccctctgac	60
acctgggaag	atggccggcc	cgtggacett	cacccttctc	tgtggtttgc	tggcagccac	120
cttgatccaa	gccaccctca	gtcccactgc	agttctcacc	ctcggcccaa	aagtcatcaa	180
agaaaagctg	acacaggagc	tgaaggacca	caacgccacc	agcatcctgc	agcagctgcc	240
gctgtctcagt	gccatgcggg	aaaagccagc	cggaggcatc	cctgtgctgg	gcagcctggt	300
gaacaccgtc	ctgaagcaca	tcctctggct	gaaggtcatc	acagctaaca	tcctccagct	360
gcagggtgaag	ccctcgggca	atgaccagga	gctgctagtc	aagatcccc	tggaacatggt	420
ggctggattc	aacacgcccc	tggatcaagac	catcgtggag	ttccacatga	cgactgaggc	480
ccaagccacc	atccgcctgg	acaccagtg	aagtggcccc	acccgcctgg	tcctcagtga	540
ctgtgccacc	agccatggga	gcctgcgcac	ccaactgctg	cataagctct	ccttcctggt	600
gaacgcctta	gctaagcagg	tcataaacct	cctagtgcga	tccatgccaa	gggtggccaa	660
ctgatcgtgc	tggaagtgtt	tccctccagt	gaagccctcc	gccctttgtt	caccctgggc	720
atcgaagcca	gctcggaagc	tcagttttac	accaaagggtg	accaacttat	actcaacttg	780
aataacatca	gctctgatcg	gatccagctg	atgaactctg	ggattggctg	gttccaacct	840
gatgttctga	aaaacatcat	cactgagatc	atccactcca	tcctgctgcc	gaaccagaat	900
ggcaaatata	gatctggggg	cccagtgctc	ttgttgaaag	ccttgggatt	cgaggcagct	960
gagtcctcac	tgaccaagga	tgcccttggtg	cttactccag	cctccttggtg	gaaaccagc	1020
tctcctgtct	cccagtgaa	acttggtggtg	cagccatcag	ggaaggctgg	gtcccagctg	1080
ggagtatggg	tgtgagctct	atagaccatc	cctytctgca	atcaataaac	acttgccctg	1140
gaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaa			1175

<210> 16

<211> 2374

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (556)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2344)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2346)

<223> n equals a,t,g, or c

<400> 16

gatcccacca	caacttaatg	ttaacatttt	aaattatttc	tttttttttc	atacatatgc	60
atagacaatt	actgggtttt	tgttttsttt	tttgtttttt	tttcaagcga	cattgtgatt	120
gtattctttt	ataccttatt	gggtttgtwt	ttttacttac	catggtaaaa	atccatttga	180
gtgagcattc	ttgagtgggt	ttgcattgtg	tcttcacaca	gttgtagcat	aatttraagct	240
gcttttggcc	ttgctctggg	aaagcagtg	agcacacact	cttaatttct	aaggaagtgt	300
agcgttccct	tgatgaatcr	ggagagagat	catgaacttg	gtgtatgagg	tacwctgggtc	360
tgtcttttca	agctggggta	agtaggcata	tgcaactacc	agtcataat	tttgtaaccc	420
actggcccg	gtgtgcccgt	attgctgagg	ctagacagat	ccatcagggc	ttagacagtt	480
agtgaagggtc	ctgggagtag	gcgaaggga	gtccagttaa	tgtgtgaagc	tgctgggaat	540
tgtgtagtgt	gcaganattt	gagcaggtaa	agcttccgat	gtggaccaag	tacaagtact	600
aagatcacgg	cacatcctct	aaaaggaagt	cagatatttg	ggactgtcca	ggaaatctgg	660
gagatagtgt	gatggtggca	tagctctttc	tcatctgaca	ctttttatgc	ttactcagag	720
tacgaatgcc	aagtattgag	acaatacaga	ataatatttg	aattgagata	ttcctagaaa	780
gacaggtcaa	catggctctt	ggtcaagaat	gacttagagt	tccttgccca	gtgctcggct	840
ggcttcatgt	catagttgag	gctgggtcca	cagggtggat	aagattccac	caaagctcac	900
caggctggta	ctgacccag	agtgtctccc	agggccctacc	ttatttttagc	attcttttga	960
attgagtttc	agaggtgatt	cagtagtaaa	tgtatgggga	gaagaattag	aaaaacccat	1020
ctctttttta	gcagtcata	ttcttaaaaa	tctgaattgt	tttacaacag	agtcagcaaa	1080
cctttctagc	atttctaaaa	gatggcagtt	tcttggagac	cacatctttg	caaggcagta	1140
ttttcaaaat	ataaaatagc	cccaaaacca	aaccttttaa	catgaagggc	aaatgggtaa	1200
agacttaata	ttctttttgt	gtcaagtata	cttaatgtaa	atctagttgc	ttgtgaagta	1260
gtaatagtaa	tagctcctat	tattttgagc	actgactata	tggtctcagc	ctctgtgctt	1320
ggatctgaca	tcactatctc	atttctttct	ctcagcagtt	ctggaaggag	tacctaccct	1380
gggtatgctt	tgcccagggt	cacagaagtt	actgaattag	agcagggtat	caccccaagt	1440
gtaaaaactc	aagagctcct	gcacttaggg	gtttacttca	ctaatactct	aacctcattt	1500
attatgctta	aggctttttc	tcagctgacc	tgactttgct	ttaggtcatt	cttttttatg	1560
ccagcactgt	ttgaaagtgc	atgtcaagcg	gctagctcca	catttgggtct	tcgaaaggga	1620
aacgcatgca	gttaaaaagt	aatgtacatg	atggaattgg	gaggatcata	gtctcagttt	1680
cccccccyct	ttctcccatc	taggagacct	ccrtggactg	cagcaaaatt	aaaaataaag	1740
cacagacaac	agaattattc	ttcactgaga	gagtttaata	tgcgtttcta	acaccatcta	1800
tacttgcttt	gttggtcttg	aggctcatca	cacacattct	ggttattcca	gagctagaag	1860
ctcttctggt	tgctaactca	gttataagaa	gatgaaagac	ataactagmc	ttacgtattt	1920
cagtagtttg	cyccttaatt	tttcccytac	ycytagtttc	caggcgacct	cccaagaagg	1980
gtatcagtcg	actggataaa	cagatgagaa	agttcacaga	tataaggaaa	aaaagcagat	2040
cygcacacgc	agcattgagg	gcaacaaaa	gccattgtga	ccttgccygg		2100
aatgtgtccc	catctctact	ctaagaaatg	cgcaattggc	tctttggaga	aagaagatat	2160
tttaaaacat	ttttagtgtg	tctgtaaatg	gttcagcgtg	tatcagatgt	tgtcatagga	2220
ctcacatttc	tctcagttat	atttaaaacc	gttggtgtact	ttgtacaaag	gaatactagt	2280
catacttcta	taaactttac	acaataaaat	ttcattctgg	twaaaaaaaa	aaaaaggggg	2340
gccncnctaa	aaaaccaagc	ttactttccc	ttgc			2374

<210> 17
<211> 1595
<212> DNA
<213> Homo sapiens

<400> 17
ggcacgagcc gggtatttta gaactagagt tgagatgaat atgacacttc ataattacac 60
tcactacatt tttcacaaat tattttttta tgctgaaaag agtaatttta cttgttgaga 120
tgttttattca ttttctaatac tatgctaaaa gcttttatca taagtcttgg gaacagttgt 180
catttacaca ttacttactg cagatatctt gactaaatca ggagggaggt gtttaatcat 240
ttgatatgta tagttgacac tcaggaatag taatgcctaa aatttacagt ggatgaggtc 300
tgtttcaagt ttaattcttt ttaaaaatgt tttacttatt tttaaatacac tttgaaaaaa 360
ttgacctcca gaatgctgtt ttatgaaatt ggcaaataaa tgaagggtatt caatttttga 420
ggaagaatga cattgccaca aaataccttt ttgtgacacc tattaaacca ctatgaaaat 480
aacttttcag gatctatttc ctatgtggaa ttctttcaaa tgctttcttc acggaaatgt 540
tttctcactc ttcgttttgt tccccttact tacatgtttc tcttttctt atactgtgaa 600
ctctggaaca aaactagatt ggggtggttg gttggttggt tgggtcttct tggagtatat 660
gtatatcaac aaaggatttg aggtcacttc tgagaataat atatcaaaaa gggtagtggt 720
tagagaaaat ataagaataa aatccagccc agagagagta ctaagaatgt aaatgtgagt 780
aaaaagccag ttttgaattc tatttacgtg acagtgatgc atgctgtcat ttaattctca 840
caaaaactct cgttatata tttatcctca tttgtagat gaggatagca ggcttagaag 900
gatttggtta tttactaatg tcatgatagg gatttaaate tgggaattgaa gtaatttgtgt 960
ctcactactc taagctaatt gaacatcaga cattaagaga gatgatcctt attgaagtaa 1020
accagttcct aaaggatcct ttagtatata atacttgga atcctgttca gagtaatttt 1080
taggaagtag gtacacagta gtaagtcttc tgttgccctc gttggccata gatggaattc 1140
tgtgtttgcc actcattaga aatttaggaa ttttaggcag ggcacggtgg ctcacacctg 1200
taatccagc actttgggag gccgaggcgg gcagatcacc tgaagtcagg agttcaagac 1260
cagcctgacc aacatgggga aaccccgctc ctactaaaaa taaaaaatta gccgggctgt 1320
gtggtgcata cctgtaatcc cagctactcg agaggccgag gcgggtgaat cacttgaggc 1380
cagaagttca agaccagcct ggccaacatg gcaaaactcc gtctctacta aaaatacaaa 1440
aattagccag gtgtggtaac gcagacctgt aatcccagct actcgaatca ggagaatcac 1500
ttgaaccag gaggcagagg ttgtcagtga gtccgagatc gcactctagc ctgggtgaca 1560
gagtgaact ctgtctaaaa aaaaaaaaaa aaaaa 1595

<210> 18
<211> 1287
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1188)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1202)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1230)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1264)
<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1277)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1282)

<223> n equals a,t,g, or c

<400> 18

aattcggcac	gagaatTTTT	tttgggtgatg	gcatgttcag	aatcttggat	ccctaagttc	60
aatatattgg	acatatTTtag	gaactctgga	aattatgttg	ttttcacata	tctagtaact	120
tactagatga	atcagtagat	ttcattaaag	tatatcta	aacagataat	tatgatgtac	180
ttctgggttg	acatgcatgt	ctctcattat	cagctatcag	tattagtgtc	atgctttgga	240
gacagttatc	ttttgaaggt	tttgggggttc	ttatgaacct	catttttccc	aggaagtttc	300
tgtaattcct	cctatgccta	ttcttgtctt	ttctatctgc	ttgcagtgtg	cgttatttag	360
atcagaggca	attatTTTTc	aggaagaaaag	aaatcatcaa	gtgacactcc	taaaggcagt	420
aaagacaaaa	tttcagtcgt	gaaccgggtct	cagaakgcct	gtattagaat	atgcaaagtc	480
catccaaatt	atatccaaat	atacttgtgg	cacagtgtc	ccagttttta	aaatgagacg	540
ttactatgta	gggcagaagt	gccaatgagg	agagagaagg	agctgttcag	tttgccctcc	600
agccgccacc	tccttctatt	attggctgaa	tgaattagt	caaaattagt	agccaaaagg	660
gtagacagt	tgaatggaag	ggaggagaag	gacagaaact	ttaatctcca	ggaaagctta	720
tttatccttt	aaaaaatgga	aagttgggca	ggcgcagtgg	ctcacgcctg	taatgccagc	780
actttgggag	gccgaggcgg	gcagatcacg	aggctcaggag	atcgagacca	tcctggctaa	840
cacagtga	ccctgtctgt	actaaaaaaa	aaaaaataga	aaaagccagg	cgtgggtggca	900
ggcgcctgta	gtcccagcta	ctcgggaggg	tgtggcagga	gaatgggtgtg	aacctgggag	960
gcgagccttg	cagtgagccg	agatcgacc	actgcactcc	agcctgggca	acagagcaag	1020
actccgcctc	aaaaaaaaaa	aaggaaagtt	gagtgtattc	catgtacctg	aacatgctat	1080
ttaaaactgt	gggctacttt	cagaatgtag	actaatgk	tctcgaccat	tggaatgaat	1140
gagaatttgk	atttgatagg	aaagtcagaa	agtcctcgag	agtccttnta	aaaccgggcc	1200
gngggcccca	tcgaatTTTT	caaccggggn	tgggggtacc	caggtaaagt	ggtaccccaa	1260
attnggcccc	tataagnnga	gncggaa				1287

<210> 19

<211> 1396

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (668)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (739)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (751)

<223> n equals a,t,g, or c

<400> 19

gctggtaacc	aggtggaacc	atttcacgtg	tccctcccca	gctgcctcag	tcccttccc	60
cacctgggcc	acagcatggg	ggttccctca	cccaccgcct	ggcctctct	tgccctgttc	120
cacactcaga	aaaaagcaag	gatcagacaa	gaagaagagt	ccccaccct	cccgctccc	180
caggagctgg	cgcttctctg	gctaaggggtg	tttttttagag	tgatgttttt	tctcctctgt	240

ctcgttgccc	tggagatcaa	agggttcact	ttctcagcga	ggggtgccag	ggacagattt	300
ctaaacaagt	ctggaccgca	gccaggaaaa	aagatgaaaa	caacacactg	taaacagcct	360
ctattcagca	aacctggtca	ggtcagaggg	gctytgagga	aagcaagagg	gaggcaggag	420
gagagggaag	cggtggggat	gtgggggggg	cgggggcaca	gttatcctga	atacataaaa	480
acaagtgagg	tactgaggt	cagggatagt	cccaaaccatc	cccaagtcca	gcctttcctg	540
acaaccaggg	ttacatgcag	agtcccaggc	catctgcagg	ttttggaggc	cctgtgcggg	600
gcctgggggt	ctatgtttta	acacgccctt	gtggtggtcc	aagtycccag	aascagggga	660
agggcgantc	tgggctctga	atggcargtg	gggcagctcc	amctcatcct	cctacatggc	720
accagcact	gggctgcang	cytgggtccc	nacttgccgc	aggaatcaat	cctgccagct	780
cagagccsc	gtgtgacaaa	caccccagga	acagaggaga	catgagaaag	ggactcacca	840
gccactgcc	caggatgtag	aagtcgtcgc	aggagaagat	ggtgccgggg	taggaggaga	900
agaccagctt	gttgccggga	accagcgggt	agtcccctgt	gcaggcagag	cgagccaggg	960
atgctgttca	gacaggcaca	ggtggaggcc	cctgcaccct	acctaacaag	acacaggcac	1020
aggggcacag	gcaggccctyc	gaggaagccc	ccactgtgtc	ctttttgtca	tttagcaaat	1080
gaggtcattg	ggcatataaa	agtgcataa	cgtgcaagta	aaaataaaag	ctagcagcaa	1140
aacttatata	gttggscccty	catgtccgtg	ggttccacat	ccttggattc	aatsgamtgg	1200
ggacaaaaaa	tactagga	aaaacatgat	taaaaagaaa	caacacagct	gggtgcagtg	1260
gytsacacct	gtaatccctg	cactttggga	ggccaaggca	ggcggatcac	gaggtcagga	1320
gaccaagacc	atcctggcta	acacgggtgaa	acccgtctct	actaaaaata	caaaaaaaaa	1380
aaaaaagggc	ggccgc					1396

<210> 20

<211> 1277

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1207)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1272)

<223> n equals a,t,g, or c

<400> 20

cgtttattca	gcagaacatc	agcttcctgc	tgggctacag	catccctgtg	ggctgtgtgg	60
gcctggcatt	tttcatcttc	ctctttgcc	ccccgctctt	catcaccaag	cccccgatgg	120
gcagccaagt	gtcctctatg	cttaagctcg	ctctccaaaa	ctgctgcccc	cagctgtggc	180
aacgacactc	ggccagagac	cgtcaatgtg	cccgctgtct	ggccgacgag	aggtctcccc	240
agccaggggc	ttccccgcaa	gaggacatcg	ccaacttcca	ggtgctgggt	aagatcttgc	300
ccgtcatggg	gacctgtgtg	ccctactgga	tgggtctactt	ccagatgcag	tccacctatg	360
tctgtcargg	tcttcacctc	cacatcccaa	acattttccc	agccaaccgg	gccaacatct	420
ctgtggccct	gagagcccag	ggcagcagct	acacgatccc	ggaagcctgg	ctcctcctgg	480
ccaatgttgt	ggtggtgctg	attctggtcc	ctctgaagga	ccgcttgatc	gaccttttac	540
tgctgcgggt	caagctgctt	ccctctgctc	tgcagaagat	ggcgtgggg	atgttctttg	600
gttttacctc	cgtcattgtg	gcaggagtcc	tggagatgga	gcgcttacac	tacatccacc	660
acaacgagac	cgtgtcccag	cagattgggg	aggtcctgta	caacgcggca	ccactgtcca	720
tctggtggca	gatccctcag	tacctgtcca	ttgggatcag	tgagatcttt	gccagcatcc	780
caggcctgga	gtttgcctac	tcagaggccc	cgcgtcccat	gcagggcgcc	atcatgggca	840
tcttctcttg	cctgtcgggg	gtgggctcac	tgttgggtc	cagcctagt	gcactgctgt	900
ccttgcggcg	gggtggtgtg	cactgcccc	aggactttgg	gaacatcaac	aattgcccga	960
tggacctcta	cttcttctctg	ctggctggca	ttcaggccgt	cacggctctc	ctatttgcct	1020
ggatcgctgg	acgctatgag	agggcgctccc	agggcccagc	ctcccacagc	cgtttcagca	1080
gggacagggg	ctgaacaggc	cctattccag	cccccttgct	tactctacc	ggacagacgg	1140
cagcagtc	agctctggtt	tccttctcgg	tttattctgt	tagaatgaaa	tggttcccat	1200
aaataanggg	catgagccct	tcctcaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1260
aaaaaaaaaa	anaaaaa					1277

<210> 21
<211> 1781
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1494)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1496)
<223> n equals a,t,g, or c

<400> 21
gctgggtgcc atggcggcag cggcgggtgac aggccagcgg ctgagaaccg cggcgggccga 60
ggaggcctcg aggcgcagc gggcgccgcc agaccactgc caggctcagg cggcggcccg 120
gctggggcgac ggcgaggacg caccgggtgcg tccgctgtgc aagccccgcg gcactctgctc 180
gcgcgcctac ttcttggtgc tgatgggtgt cgtgcacctg tacctgggta acgtgctggc 240
gctgctgctc ttctgtcact acagcaacgg cgacgaaagc agcgatcccg ggcccccaaca 300
ccgtgcccag ggccccgggc ccgagcccac cttaggtccc ctcaccggc tggagggcat 360
caagggtggg cagagcgta aggtccagct ggtcaccgac agggatcact tcatccgaac 420
cctcagcctc aagccgctgc tcttcgaaat ccccggttc ctgactgatg aagagtgtcg 480
gctcatcatc catctggcgc agatgaagg gttacagcgc ascagatcct gcctactgaa 540
gagtatgaag aggcaatgag cactatgcag gtcagccagc tggacctctt ccggctgctg 600
gaccagaacc gtgatgggca ccttcagctc cgtgaggttc tggcccagac tcgcctggga 660
aatggatggt ggatgactcc agagagcatt caggagatgt acgccgcgat caaggctgac 720
cctgatggtg acggagtgtc gagtctgcag gagtctcca acatggacct tcgggacttc 780
cacaagtaca tgaggagcca caaggcagag tccagtgagc tgggtgcgga cagccaccat 840
acctggctct accaggggtga ggggtgccac cacatcatgc gtgccatccg ccagaggggtg 900
ctgcgcctca ctgcctgtc gcctgagatc gtggagctca gcgagccgct gcagggtgtt 960
cgatatggtg agggggggcca ctaccatgcc cactgggaca gtgggcctgt gtaccagag 1020
accatctgct cccataccaa gctggtagcc aacgagctgt tacccttcga gacctcctgc 1080
cgctacatga cagtgtgtt ttatttgaac aacgtcactg gtggggggcga gactgttttc 1140
cctgtagcag ataacagaac ctacgatgaa atgagtctga ttcaggatga cgtggacctc 1200
cgtgacacac ggaggcactg tgacaaggga aacctgcgtg tcaagcccca acagggcaca 1260
gcagtcttct ggtacaacta cctgcctgat gggcaagggt ggggtgggtga cgtagacgac 1320
tactcgctgc acgggggctg cctggtcacg cgcggcacca agtggattgc caacaactgg 1380
attaatgtgg accccagccg agcgcggcaa gcgctgttcc aacaggagat ggcccgcctt 1440
gcccgagaag ggggcaccga ctcacagccc gagtgggctc tggaccgggc ctancncgat 1500
gcgcgcgtgg aactctgagg gaagagttag ccccggttcc cagccgcggg tcgccagttg 1560
cccaagatca ggggtccggc tgccttctg tcctgtgca gactaaagg ctggccaatg 1620
tcttgcccca ccccgccagc cgcgatacgg cgcagttcct atattcatgt tattttattgt 1680
gtactgactc catctgcccc gtcaataaaa aaaccacaag gttcgaaaaa aaaaaaaaaa 1740
aaaattgggg gggggggccc gtaaccatt tggcccttta g 1781

<210> 22
<211> 1491
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1425)
<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1426)

<223> n equals a,t,g, or c

<400> 22

ggaagtgtag	gtacagattc	aggatcataga	aaggcccata	tgtcccaata	ttatttytyt	60
tacagagaaa	acagatgaaa	ttracaattt	tttttktytt	tccacagacc	atcaccggcc	120
tcctgcagak	tttgatgtcc	aggcaagtag	aggatgtggc	tttccttccc	cttcctcatc	180
ccgtcttctc	tttttccttt	ttctttccat	tggtttaagt	agatcattgt	gcaaacattg	240
cgggcaaggg	gagagaaggc	agtgggtctc	gctaggtcct	caccctcagc	tgctgtctcc	300
agacagaagc	tgacttgagt	gctggccaga	gaatcaggaa	ccagtcactt	cccacgaaag	360
caggacagag	acaccgcctg	tttcagtccc	aaagagctgg	ccagaacggt	ggggcggtgtg	420
tggggaaaat	ggtgtgatgc	tgggctggct	tgggaccctg	tgaggggagg	gaacggaagc	480
aaaagcttga	tgttcttggc	tccacttggg	aaactgaaac	cagccttctc	tgtggtttac	540
tggcccagaa	ctgaagtatc	tccattccca	ctcaggattg	caggggtggg	caggggggtgc	600
agctcagctg	tgtccaagaa	taggctctca	ggagaacggg	ctgggctctc	tcagccgagg	660
gctgggtcag	aacttcagga	aatctaagtc	ctggactcgt	tagccccaag	ctgggggtggg	720
tcccatgctg	tgctggcctc	tggagggtgg	ggcraaaagc	ctgagcatat	ggcgagcttg	780
tggctctcct	gaaggaggag	aggcgactc	ctggaccagc	cagagagcca	gaggtggctg	840
acggtgttct	cacctgaacs	gytacctatg	tgttgacacg	gtgtccccag	tttgggacct	900
attctgcact	ctgcatctca	ctcctctatt	ctacctcctt	acctcaaaac	aactccgctt	960
ttcctcttga	atctctgatt	tcatgaatgc	tcccccatcc	atcccagagg	gaggaatctt	1020
gcactgtctt	tggaaagaaa	ggaaaggatg	ggagaggggc	tggaaagcatt	ggctagtaga	1080
tgatcgaggc	tttctatacc	accaccaagg	tgtgtatttt	ctcatcatcc	tggttgtgtc	1140
aggctgtccc	tagagatcac	cctgacctca	gtgtttaaga	agaagggcca	gatgcagtgg	1200
ctcatgcctg	taatcctagc	actttggaag	gccgaggcag	gcagatcacc	tgaggtcagg	1260
agctcgaggc	cagcctggcc	aacaaggaga	aaccctgtct	ctactaaaaa	tacaaaaatt	1320
agccaggcat	ggtggttaagc	acctataaacc	ccagctactt	gggggctgag	acaggagaat	1380
tgcttgaacc	tgggaggcag	aggttgcagt	gagccgagac	cacgnngttg	cactccagcc	1440
tgggcaacaa	gagtgaact	ctgtctcaaa	aaaaaaaaaa	aaaaactcga	g	1491

<210> 23

<211> 1839

<212> DNA

<213> Homo sapiens

<400> 23

aattcggcac	gagtgcaggt	cgactctaga	ggatccccgc	taagaagcta	gggctattgg	60
tcttcccata	cacacatcag	aactgaggca	ccatgcaagg	gggccagaga	cctcatctcc	120
tcttgctgct	gttggctgtc	tgcctggggg	cccagagccg	caaccaagag	gagcgtctgc	180
ttgcggacct	gatgcgaaac	tacgaccccc	acctgcggcc	ggctgagcgc	gactcagatg	240
tggtaaatgt	cagcctgaag	cttaccttga	ccaacctcat	ctccctgaat	gaacgagagg	300
aggccctcac	aactaacgtc	tggatagaga	tgcaatgggtg	cgactatcgc	ctgcgctggg	360
acccaaaaga	ctacgaaggc	ctgtggatat	tgaggggtcc	atctactatg	gtctggcggc	420
cagatatcgt	cctggagaac	aatgtggacg	gtgtcttcga	ggtggctctc	tactgcaatg	480
tcctcgtgtc	ccccgacggg	tgtatctact	ggctgccgcc	tgccatcttc	cgctcctcct	540
gctccatctc	tgtcacctac	ttccccctcg	attggcagaa	ctgttccctc	atcttccaat	600
cccagactta	cagcaccagt	gagatcaact	tgacgtgag	ccaggargat	gggcaagcca	660
ttgagtggat	cttcattgac	ccggaggctt	tcacagagaa	tgggragtgg	sccatccggc	720
accgaccggy	taaaatgctc	ctggactccg	tggctctctc	agagraggcg	ggccaccaga	780
aggtgggtgt	ctacctgctt	atccagcgca	agccctctct	ctacgtcatc	aacatcatcg	840
ccccctgtgt	gctcatctcc	tcagtgcgca	tcctcatcta	cttccttccct	gctaaggcgg	900
gcggccagaa	atgcacagtg	gccaccaacg	tgtctctggc	ccagactgtc	ttccttttcc	960
ttgtggctaa	gaagggtgct	gagacctccc	aggcagtgcc	actcatcagc	aagtacctga	1020
ccttctcat	ggtggtgacc	atcctcatcg	tcgtgaaact	tgtgggtcgtg	ctcaatgtgt	1080
ccttgccgtc	ccccacaca	cactccatgg	cccggtgggt	ccgcaagggtg	ttcctgaggg	1140
tcctgcccca	gctgttacgg	atgcatgtgc	gccactagc	tccagctgct	gtccaggatg	1200
cccggttccg	actccagaat	ggctcttcc	cagggtggcc	catcatggct	cgagaggaag	1260
gggacctctg	tctgcctcga	agcgaactcc	tctttagcca	aaggcagcgc	aatggattag	1320

12

tgacggcagt	attggagaag	ctagagaatg	gtccagaagt	gaggcagagc	caggagtctt	1380
gtggcagcct	gaagcaagcc	tccccagcca	tccaggcctg	tgtggatgcc	tgtaacctca	1440
tggctcgtgc	ccgacgccag	cagagtcact	ttgacagtgg	gaacgaggag	tggttgctgg	1500
tgggcccag	gctggaccga	gtctgcttcc	tagccatgct	ctccctcttc	atctgtggca	1560
ctgctggcat	cttcctcatg	gcccactaca	accaagtgcc	tgacctgccg	ttccccggag	1620
acccccgccc	ctacctgcct	ttgccagact	gagccaacca	atccctcctg	ggccctggag	1680
tcagctatga	gggccatgct	gtttgtagag	ctgtatcccc	tggtgatgct	gagtgtgctc	1740
ttggggaaat	acccaaggct	tcctgggaga	agatagagaa	ataaagagac	agaggggaaa	1800
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aactcgtag			1839

<210> 24

<211> 1384

<212> DNA

<213> Homo sapiens

<400> 24

ggcacgagca	gttattttca	aaatggctat	ggaaaacacg	taagttttta	aatatgccct	60
ctttctcggt	ttaaaaaatt	attactattg	tccatacatg	ttactctttt	catctagatt	120
tatcatgttt	ctttggcctc	cagtctctgg	tgtttgccca	agctttatta	gagacaggtc	180
atttctacct	atgtgtcatt	ttatctatgt	cttgatctta	tgtaattcaa	ttgctcttta	240
agattatggt	ctcttctcat	gtttggttta	tccattatcc	aaattttcca	tttcttttaac	300
ctgttatccc	ttgactcttt	acagttctac	ctttttattc	acttagtctt	ttaccctttt	360
tttattcggt	cacccctttt	tggtgtttca	ggtactcctt	acttatctcc	ttagcctttt	420
cttcttcatc	ttctttctta	cttttctcct	acttctcatt	ttacataata	cttacttttt	480
gcttcagctc	tcaaccattg	tcaatcttgt	ttttccttat	attccatttt	actttctgaa	540
ctactcttta	atctcctggt	caacactacc	tttccttctt	ttttatcccc	tcttattttac	600
acggtgatta	caacagtttg	gtatagtctg	atttatctga	ttgtaaaatt	gatgagttgg	660
atgtaccaaa	aatataagga	agctaaattc	aaagaaggta	aaagatttgc	ttgtgtcacc	720
tagctgggta	attttgccat	atgcattggt	tctctacata	gtctatgtag	tcaaacagggt	780
ttcatttaga	aatcattccc	cataagaagg	gtttcaattt	gatttgaaca	ggcagagatg	840
gaaaaaattt	cctctctgat	aactactgct	actgttgtat	accagtagaa	atataacagc	900
agcacttagg	ttagaagaag	ctcattagct	attcagaata	aatttcatth	ttcttaattt	960
ttggtaatca	tatctcagcc	tggtgaattt	aacttaaaact	ctgaaagaat	tttggttgcc	1020
atttaatttt	taggtttcct	taatgatagg	gacctaatat	tttggtttta	aaaatttgtc	1080
ttggttgagg	gcagtggtc	atgcctgtaa	tccagcact	ttaggaagcc	aacattggag	1140
gattgcatga	gcccaggatt	tcgagaccag	cctgggcaac	acagtgaac	ctcatctcta	1200
caaaaagtta	aaaaattaac	caactgtggt	gccacatgcc	tgtagtccca	gctgcttggg	1260
aggatgaggt	gagaggattt	cttgagtcca	ggagtttgag	gctgcagtga	gctatgatca	1320
cactcctgct	cttcagccta	ggtgacacag	caggacacta	tctttgaaaa	aaaaaaaaaa	1380
aaaa						1384

<210> 25

<211> 1681

<212> DNA

<213> Homo sapiens

<400> 25

tctctggggc	aatatggcag	cgcccagcaa	caagacagag	ctggcctgga	gtccgcggct	60
ggccgcgtga	gtaggtgatt	gtctgacaag	cagaggcatg	agctgggtcc	aggccaccct	120
actggcccga	ggcctctgta	gggcctgggg	aggcacctgc	ggggccgccc	tcacaggaac	180
ctccatctct	caggctccctc	gccggctccc	tcggggcctc	cactgcagcg	cagctgcccc	240
tagctctgaa	cagtcctctg	ttcccagccc	accggaaccc	cggcagaggc	ccaccaaggc	300
tctggtgccc	tttgaggacc	tggttgggca	ggcgcctggt	ggggaacggg	acaaggcgag	360
cttctgagc	acggtgcaga	aatgtgcgga	gcacagcgtg	cgtaagcggg	gccacattga	420
cttcatctac	ctggccctgc	gcaagatgcg	ggagtatggt	gtcgagcggg	acctggctgt	480
gtacaaccag	ctgctcaaca	tcttccccaa	ggaggtcttc	cggcctcgca	acatcatcca	540
gcgcactctc	gtccactacc	ctcggcagca	ggagtgtggg	attgctgtcc	tgagagagat	600
ggagaaccac	ggtgtgatgc	ccaacaagga	gacggagttc	ctgctgatcc	agatctttgg	660
acgcaaaaagc	taccccatgc	tcaagttggt	gcgcctgaag	ctgtggttcc	ctcgattcat	720

13

gaacgtcaac	cccttcccag	tgccccggga	cctgccccag	gaccctgtgg	agctggccat	780
gtttggcctg	cggcacatgg	agcctgacct	tagtgccagg	gtcaccatct	accaggttcc	840
tttggccaaa	gactcaacag	gtgcagcaga	tcccccccag	ccccacatcg	taggaatcca	900
gagtcccgat	cagcaggccg	ccctggcccc	ccacaatcca	gcccggcctg	tctttgttga	960
gggccccttc	tccctgtggc	tccgcaacaa	gtgtgtgtat	taccacatcc	tcagagctga	1020
cttgctgccc	cgggaggaga	gggaagtggg	agagacgccg	gaggagtggg	acctctacta	1080
cccgatgcag	ctggacctgg	agtatgtgag	gagtggctgg	gacaactacg	agtttgacat	1140
caatgaagtg	gaggaaggcc	ctgtcttcgc	catgtgcatg	gcgggtgctc	atgaccaggc	1200
gacgatggct	aagtggatcc	agggcctgca	ggagaccaac	ccaaccctgg	cccagatccc	1260
cgtggctctc	cgcctcgccg	ggtccacccg	ggagctccag	acatcctctg	cagggtctga	1320
ggagccgccc	ctgcccagag	accaccagga	agaagacgac	aacctgcagc	gacagcagca	1380
gggcccagagc	tagtctgagc	cggcgcgagg	gcacgggctg	tggcccagag	aggcgggtga	1440
ctgaaggcat	gagatgccct	ttgagtgtac	agcaaatcaa	tgttttcctg	cttggggctc	1500
tcttccctca	tctctagcag	tatggcatcc	cctccccagg	atctcgggct	gccagcgatg	1560
ggcaggcgag	acccctccag	aatctgcagg	cgcctctggg	tctccgaatt	caaataaaaa	1620
ggggcgggag	cgctgttggg	tgtgcgcaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1680
a						1681

<210> 26

<211> 1949

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1130)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1948)

<223> n equals a,t,g, or c

<400> 26

gctacgccgt	gggcacttcc	tcaacgacct	gtgcgcgtcc	atgtggttca	cctacctgct	60
gctctacctg	cactcgggtg	gcgcctacag	ctcccgccgc	gcgggctgct	gctgctgctg	120
ggccagggtg	cgacgggctg	tgcacaccgc	tcgtgggcta	cgaggccgac	cgcgccgccca	180
gctgtctgcg	ccgctacggc	ccgcgcaagg	cctggcacct	ggtcggcacc	gtctgcgtcc	240
tgtgtctcct	ccccttcac	ttcagccctc	gcctgggctg	tggggcgggc	acgccgagtg	300
ggctgccctc	ctctactacg	gcccgttcat	cgtgatcttc	cagtttggtg	gggcctccac	360
acagatctcc	cacctcagcc	tcaccccgga	gctcgtcacc	aacgaccatg	agaagggtga	420
gctcacggca	ctcaggtatg	cgttcaccgt	ggtggccaac	atcaccgtct	acggcgccgc	480
ctggctcctg	ctgcacctgc	agggctcgtc	gcgggtggag	cccacccaag	acatcagcat	540
cagcgaccag	ctggggggcc	aggacgtgcc	cgtgttccgg	aacctgtccc	tgctggtggt	600
gggtgtcggc	gccgtgttct	cactgctatt	ccacctgggc	acccgggaga	ggcgccggcc	660
gcatgcggag	gagccaggcg	agcacacccc	cctgttggcc	cctgccacgg	cccagcccct	720
gctgtctctg	aagcactggc	tccgggagcc	ggctttctac	cagggtggga	tactgtacat	780
gaccaccagg	ctcatcgtga	acctgtccca	gacctacatg	gcatgtacc	tcacctactc	840
gctccacctg	cccaagaagt	tcacgcgcag	cattcccctg	gtgatgtacc	tcagcggtt	900
cttgtcctcc	ttcctcatga	agcccatcaa	caagtgcatt	gggaggaaca	tgacctactt	960
ctcaggcctc	ctggtgatcc	tggcctttgc	cgctgggtg	gcgctggcgg	agggactggg	1020
tgtggccgtg	taygcagcgg	ctgtgctgct	gggtgctggc	tgtgccacca	tcctcgtcac	1080
ctcgtggtcc	atgacggccg	acctcatcgg	tccccacacg	aacagcggan	ckttcgtgta	1140
cggctccatg	agcttcttgg	ataaggtggc	caatgggctg	gcagtcattg	ccatccagag	1200
cctgcaccct	tggccctcag	agctctgctg	cagggcctgc	gtgagctttt	accactgggc	1260
gatggtggct	gtgacggggc	gcgtgggctg	ggccgctgcc	ctgtgtctct	gtagcctcct	1320
gctgtggccg	acccgcctgc	gacgctggga	ccgtgatgcc	cgccctgac	tcctgacagc	1380
ctcctgcacc	tgtgcaaggg	aactgtgggg	acgcacgagg	atgcccccca	gggccttggg	1440
gaaaagcccc	cactgcccct	cactcttctc	tggacccccca	ccctccatcc	tcaccagct	1500

14

ccccgggggtg	gggtcgggtg	agggcagcag	ggatgcccgc	cagggacttg	caaggacccc	1560
ctgggtttttg	aggggtgtccc	attctcaact	ctaattccatc	ccagccctct	ggaggatttg	1620
gggtgccccct	ctcggcaggg	aacaggaagt	aggaatccca	gaagggtctg	ggggaaccct	1680
aaccctgagc	tcagtccagt	tcacccctca	cctccagcct	gggggtctcc	agacactgcc	1740
agggccccct	caggacggct	ggagcctgga	ggagacagcc	acgggggtgg	gggctgggcc	1800
tggacccccc	cgtggtgggc	agcagggtctg	cccggcaggc	ttggtggact	ctgctggcag	1860
caaataaaga	gatgacggca	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1920
aaaaaaaaaa	aggggggggg	gctagttnnt				1949

<210> 27

<211> 2286

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (2262)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2264)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2272)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2278)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2279)

<223> n equals a,t,g, or c

<400> 27

gctgatgtcg	aggttcatcc	tgaaccacct	ggtgctggcc	attccactga	gggtgctggg	60
ggttctgtgg	gccttcgtct	tgggcctatc	cagggtcatg	ctggggcggc	acaatgtcac	120
cgacgtagct	tttggttttt	ttctgggcta	catgcagtac	agcatcgtgg	actattgctg	180
gctctcacc	cataatgctc	cggtcctctt	tttactgtgg	agtcaacgat	gacaccatct	240
cattgattat	ggcaccagga	agtctgaagg	tttccacatt	cgatgatgtc	aacctaaacc	300
agcagccatc	ccgcttgctc	ctcttaggca	tttcaggctt	cctttgggat	ttcagggtgc	360
ccatgatctt	gatgtgctgc	taggctggag	cacacactgg	ccattactga	acacagccat	420
attagggaaa	gcaaaaaaac	ccaaaaaatc	ctctattgta	tattttattca	acaactgttt	480
atgtttccag	gacaactgca	aagaaaaaca	gctgaggtgg	ttatactgtt	gctgttaaaa	540
gttggtatca	gtaagatttg	tgttttgtga	taatccctaa	atcaacatac	cacttgtaaa	600
ctgaacttcg	agaaagaaac	atgatgttca	ttctgtaaat	atacatgcag	acagggtcatg	660
tactaatcct	agtccttttc	ctgaggtaga	ttttaaacag	tattttttaa	gtccaagaca	720
taggtttttc	tagttttatt	cctgaagatc	tgttgccaca	gttgggagat	ttcttcttaa	780
tcctgatatt	cttggttaagc	ttttttactt	tattatctct	ataatttatt	atctctatcc	840
atatttggtg	atcgggtagt	gggaaaagag	attataatac	ttgtctttct	ctcctctccc	900
tccatccctc	aaaagatctt	tatgcatttc	ccactactcc	cttactgtct	tttagcattc	960
agagaaaaag	ccaacttgct	taaagaggaa	tcacttaaaa	ggtaggcata	tctaagatgc	1020
tcatagaaga	ggaagaatgg	gacatggccc	catgcttatt	tttgtttaca	acgtaacatg	1080
gcatgagaga	gggcagagaa	actaagttgc	tggggaaagt	tagagggaact	gaaagtttgg	1140

15

gaataggctg	accacatatt	atgccagtga	ccagtatgac	aggagatggg	gccctgctgc	1200
cagtcattctc	cactgaataa	agaataatgc	tcctctttca	gggtaataaa	gtggggaaaa	1260
ggaacgtctt	ctcaatgcaa	gaacataagc	tttctcgtat	atacctgtat	gctacagttt	1320
ttcacatgga	attccgtttt	ctgaggtaca	gcacatttta	ggtaacagta	tttaacttga	1380
aattcatcat	gggagtctgc	tgctatacca	ggcacaagat	aaaactccaa	aatttctgtt	1440
tacattgacc	tttacattta	aagctgttca	tccatgggtgc	ctccccaaat	cataagacca	1500
aagaccacca	aacgcagggt	ggactctgct	cattattctt	tgaccagaa	agactggaga	1560
aggatatgtc	tttaagtgtc	gctctacctg	aaaagaaatc	ctttaaatga	cctatggaag	1620
tgatgtcctc	agataatctt	aatgactatt	ttggcattta	taaatagaaa	tgattatgga	1680
ctttgatctg	ccatacggag	gttcggaacc	tggagaaygg	ctgtgataag	taggttttga	1740
ttgagtga	gcatgagctt	gttcagagt	aggggcatag	tgaaaaagga	acagccatgc	1800
ctcawaatca	aatcatttgc	rttcccacag	catacctgaat	accgactacc	tcttccactg	1860
ctaaagcagc	taaactgtga	agctctaagt	ggtttgggtt	tggtgtttta	ccttagcgag	1920
atcctttaac	tgacgcaata	ttcaagccag	atatttggaa	gcaaatgata	tttccctctg	1980
cagtgtccac	aatcttgaat	attaggggca	tgaaattagg	cttaccatct	gatttgtaat	2040
tacaattttg	gaattctctg	ttttagttgc	tgaggcctga	gttttctggc	tcttaaagca	2100
tagatcattt	cacctgatgt	ttttgaagca	tcctaagtac	agtagagtag	aaaactgatt	2160
tctttgttaa	ttgtacactg	aataatgcct	tttaaaaatc	aaaataaaat	taacaaataa	2220
tggtgaaaaa	aaaaaaaaaa	aaaaaaaaa	cgaggggggg	cncnaaaaca	antcgacnna	2280
tagtga						2286

<210> 28

<211> 530

<212> DNA

<213> Homo sapiens

<400> 28

ccacgcgtcc	gaaaattttt	tcaatatatt	ttattaatct	ttttataaaa	tgaaaagaaa	60
ctcctatgat	cgattaagga	aggtgggtat	ggctgggtgg	ttcaggggtt	tttttgggtt	120
tctttttttt	tttctttgtc	tttttaacct	taagctgttt	aagttgaagc	attctcagat	180
gtttgggggg	aaacatcctc	ttaaaatggg	tccttgtgct	tgcttctctg	ggagggcggtc	240
ctgagcaggt	gaatcataag	gcatttatgc	atatgttata	tgcggaactgc	acccacctct	300
cccccccagc	ctttgcctct	tggtgtgttg	tgctgctttc	cccttacttt	gctacatttc	360
tatagtttaag	ttggtttttac	ttgaatgatt	catgtttagg	gggaaaatga	aaatctccct	420
taaaatttgt	ttcaactcct	cctgcaataa	aaataaatga	agtggcagat	gtaaaaaaaa	480
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa		530

<210> 29

<211> 1296

<212> DNA

<213> Homo sapiens

<400> 29

cttcatcagc	tgcgacctcc	tcaccgcttt	cctcttatac	cgctgctgc	tgctgaaggg	60
gctggggcgc	cgccaggttg	tggtactgt	gtcttttggc	ttcttaacct	cctgcctatg	120
gcaktatcca	gccgcggtaa	tgcggaactc	attgtgcct	ccctggtcct	gatggctcct	180
tacttgataa	agaaaagact	cgctgcgtgt	gcagctgtat	tctatgggtt	cgygggtgcat	240
atgaagatat	atccagtgc	ttacatcctt	cccataacct	tccacctgct	tccagatcgc	300
gacaatgaca	aaagcctccg	tcaattccgg	tacactttcc	aggcttgttt	gtacgagctc	360
ctgaaaaagc	tgtgtaatcg	ggctgtgctg	ctgtttgtag	cagttgctgg	actcacgttt	420
tttgccctga	gctttgggtt	ttactatgag	tacggctggg	aatttttggg	acacacctac	480
ttttatcacc	tgactaggcg	ggatatccgt	cacaactttt	ctccgtactt	ctacatgctg	540
tatttgactg	cagagagcaa	gtggagtgtt	tccctgggaa	ttgctgcatt	cctgccacag	600
ctcatcttgc	tttcagctgt	gtotttcgcc	tattacagag	acctcgtttt	ttgttgtttt	660
cttcatacgt	ccatttttgt	gacttttaac	aaagtctgca	cctcccagta	ctttctttgg	720
gtacctctgg	cttactgcct	cttgtgatgc	cactagtcag	aatgccttgg	aaaagagctg	780
tagttctcct	aatgttatgg	tttatagggc	aggccatgtg	gctggctcct	gcctatgttc	840
tagagtttca	aggaaagaac	acctttctgt	ttatttgggt	agctggtttg	ttctttcttc	900
ttatcaattg	ttccatcctg	attcaaatta	tttcccatta	caaagaagaa	cccctgacag	960

16

agagaatcaa	atatgactag	tgtatgttcc	acaccctctg	ctactgtgtt	acattctgat	1020
tgtcttgtat	ggaccagaag	agagcttttg	gacatttttt	ctgaacattc	taagcattct	1080
agtgaaggtt	cccatgttcc	aacagaactt	aaaagcaatg	tttgccttat	atataaaagg	1140
gacacaataa	ttgaggtcca	ccttctagga	aatcctagga	ctcgtttatt	tgggacatgg	1200
tgggaataaa	ggtcacatat	tggaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1260
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaa			1296

<210> 30

<211> 1979

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (968)

<223> n equals a,t,g, or c

<400> 30

gctttgccag	ggctgagccg	ggctgcctgg	tgccttcacc	gcccccgcca	wacaccacca	60
tgcwgtactc	cggcctgcgg	aactcgtagt	gcagcccctg	tcgcctcccc	ggcccctgct	120
atcccacgca	ggactggctt	cggccgcggg	ggccagcagc	ttgcracgtg	tccctgggga	180
ggcgggaatcg	ctgtgcgccc	tgagcccggg	ctcagccctt	cgctttccag	ctgcgtcctg	240
ctcccggccg	scagaggagc	ccagtggcga	tgagggcact	gctggcgctt	tgccttctcc	300
ttggctggct	gcgctggggc	ccggcggggc	cccagcagtc	cggagagtag	tgccacggct	360
gggtggacgt	gcagggcaac	taccacgagg	gcttccagt	cccagaggac	ttcgacacgc	420
tggacgctac	catctgctgc	ggctcctgcg	cgctccgcta	ctgttgcgcc	gcggccgacg	480
ccaggctgga	gcagggcggc	tgaccaacg	accgccgcga	actggagcac	ccaggcatca	540
ctgcgcagcc	tgtctacgtc	ccctttctca	tcgtcggctc	catcttcatt	gcgttcatca	600
tcctgggctc	tgtagtggct	atttattgtt	gcacctgtt	gagacccaag	gagccctcgc	660
agcagccaat	ccgcttctca	ctccgcagct	atcagacaga	gacctgccc	atgatcctga	720
cctccaccag	ccccagggca	ccctcccggc	agtccagcac	agccacgagc	tycagcttca	780
caggcggcty	catccgcagg	ttcttctcag	ccatctgggt	tcctgggtgc	accccagtat	840
ttcgcttacc	cccttcagca	gragcccca	ctggctggga	agagctgtcc	agactttcag	900
ttccagktga	cacgcccagg	ccatgaatyc	acaactcagt	cagatggcag	acaggtggag	960
ccctgctncc	attgccacat	gcaattctga	gaaaatttcc	cttgtaactg	atcagtgtcw	1020
tggaggagca	tgctaggaaa	acacagcacc	ttctaatttg	aaagttcctg	tctccaatca	1080
cagaaaggct	aaaccagaga	actgtttctg	gttttgcaaa	catgtgatca	ttacatttca	1140
atctatgcta	cttttattca	aaatatgcag	cagtttgact	ttaaagttgc	aaactggcta	1200
aaaacgtttt	actggacatt	cagctatatt	gcttagaaaa	gggtacatg	tttctttttc	1260
atataagttg	ttcattgagt	tatgatagga	atatattcat	aaataagcaa	agaaaaatac	1320
ctaattgtaa	ttatcaaagg	ttcacttaaa	aaattaacta	ttaggtaaac	ttaagggggc	1380
agtgaaaaat	ctatttatga	tttcgggagt	aacctaacca	tgaataatat	tagcatwatg	1440
agamcatttm	ctttttaaat	aaatamctaa	atttkgttta	caaymggagt	tttyccagaa	1500
tacaaggtty	caataatcac	atgaggagtt	taaagtttta	aatatatact	cagacattca	1560
ttgtaacaca	gagtgtatgt	aaaatcattt	ccccactca	ctggagggag	tattttattgc	1620
agactttttg	ttcagcaaca	tttagtgttt	cagtgaaggt	tggacagttg	gggcttaaaa	1680
cattttattg	taaaatgagc	tatgttcaaa	tgtaaatatt	tgtaatttaa	tgtatttacc	1740
mcattgactg	tactaattat	ttagtagtca	tactgttaatt	tttatgttaa	taataactgg	1800
agttcaaagt	ctagctattg	gtataatcat	ctaataattat	atatactctc	agtgcctctg	1860
aattttatgt	ttgatgacta	tataattggg	catatatctt	gttggattag	aataaataaa	1920
acactttata	ttttcatgaa	ctctaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1979

<210> 31

<211> 1274

<212> DNA

<213> Homo sapiens

<400> 31

gcccacgcgt	ccgctgttgc	tcaaaggaaa	taggagttgg	tgtgcttgtg	accaaggggt	60
------------	------------	------------	------------	------------	------------	----

17

tacacttmca	gcttttaaaa	ttctccttta	catgtgctca	gtgttttgkt	ttgtgttttg	120
gtttctgttt	tttattttta	ttcccacatt	gggcacaaga	atcagaatat	ggatagctag	180
tttaagaaac	ttttgtgggt	gcactgtagc	atagatgaca	gaatttgatg	ttccccccat	240
ctccaattca	gttcagggca	ttccacagtt	aaacagaaat	gggaacgtgg	ggctcttata	300
aatgaatggg	cgctcacagt	tttggttttc	agctcttcat	gtctgtaagt	gtgctttggg	360
graggctatg	tctgtatggt	cgattctcag	ttatcacatt	tgcctctcct	cccactacct	420
tcatgamcat	tcagtgtgtg	tcgcactgca	gtagagaga	agggacggac	agttgggtgac	480
actcagccac	attgctactt	ttatctgttc	tggtagaag	ttagatagat	ggtagattga	540
agcaattggg	tagaattagt	tgggggaata	tttatgagtt	gctgtgtttg	ttgattagtt	600
ccatctcttt	cccattttta	ctgagaattg	attatatata	gctctaagta	tataggtatt	660
taaacaaccc	cacaagcggc	tgtatcagta	acatttatta	attccactat	agtgagggag	720
gatttccatt	ctaaatacct	tattttgagg	gatttataaa	acttagttgt	aaaagagaaa	780
gcccacatag	tgggaataaa	ttgcttcagc	catttttagt	atttgagagc	actagggaaag	840
atgttttagta	gctgtgtgga	tgcctttttt	cacaccctgt	ctattgaatg	ctgcatccat	900
tcacgaagtt	aaatgttaca	tgcagttagt	ccttaatgtg	gactggatct	gtacttttgt	960
tttgatttaa	aacattttaa	gatttttgaa	gtgcagctac	tccccacgtg	catttgmtac	1020
acataaaaagt	catactgtgt	gtgcacaaag	agtacatgga	ttttccagca	taytgcttta	1080
aaaaattata	taaaactgta	aaatattaac	acctcaggct	acctgctgta	ttctgtccca	1140
ttgaccctcg	gaattggatt	tactgcaagt	gattgataat	tcaattatgt	ggcttttccc	1200
ctttaatctt	gccattttaa	ttacagtaga	aagacaaaat	caagtaaaat	aaagtgttag	1260
ataatagaaa	gagt					1274

<210> 32

<211> 1531

<212> DNA

<213> Homo sapiens

<400> 32

tcaaaaagact	acttagtgac	actataaaag	ttacgcctgc	aggtaccggg	ccggaattcg	60
cggccgcgctc	gacgaagtgc	tgaccaattg	ccactggaca	tacttgaaac	aaaataggaa	120
aatggcagca	aactcttcag	gacaagcatt	gcactctcga	gacctctct	taataaggac	180
ttccgggac	acgctgagca	gcagcatatt	gcagcccaac	agaaggcagc	tttgacgcat	240
gctcatgcac	attcatctgg	atacttcac	actcaagact	ctgcatttgg	gaaccttatt	300
cttctctgttt	tacctcgct	tgaccagaa	tgaagaaaac	atttgcgatg	gaaaagtgc	360
tttgtaatat	caaagccaa	agctactatc	attcagtgct	acatgaactg	tgactttaag	420
aattttggtg	aactttgata	ttttttgttt	gtctgaaaga	aaggaatgtg	taagtgaag	480
ctgaaagaag	aataaccagg	atgatgagag	ctgtggaagc	tgtatcgccc	aaggaattga	540
ttatgtaccg	tgactgtaac	ttttttgtaa	tgctgtttta	ctctcaatca	gactgtgaac	600
tggtggtgca	cgaagtcatt	ccccaaactcc	tagcaagttt	gactgaatat	atcatgtcca	660
cagtagatgt	tcaagaatca	tttatagtac	ttaaacttta	agaaacaagg	ctgcttttaa	720
aaaatgaact	aataggctta	aatcaattgc	atccatattt	gctgtttata	ggattgctat	780
cagtatacct	tttgcgttta	tagtcaacat	gtatcatcct	gaaatattct	ttctggactt	840
ataactactt	cccccttttt	cacttttaaa	caaacctcaa	gaataaatta	ctaaccagtc	900
ttaaccatct	tttataaaca	tatgctctta	taaatgttgt	gactagatgc	aattaaaaat	960
aatagggaat	gtggtaggtt	tttaatttgt	acatcctctt	atttagtggt	accacataaa	1020
tgatgagttt	gtgtggttct	gttttccatt	tttgttctaa	ctgaaaactt	tttggtgtgt	1080
cttgaactct	tgacctcaag	cagtcctctc	gatcctccca	ccttgccctc	ctaaagtgt	1140
gagattacag	gtgtgagcca	ctgcatctgg	cttacttatt	ttgtctattg	tctgttccac	1200
tagtatgtaa	agtccttagag	agcaagaatt	tttgtttatt	tctttctctt	cctcctttcc	1260
tttcttctct	ttttacttcg	ttcactactg	tattccacat	aaaatatatt	tggcatatag	1320
taggtgttca	atatgttgaa	ggaatgaaag	aatttataga	cttgagttgc	aatataaaat	1380
gtattttttt	ttactgtgag	ttatggcaaa	aaaagttttg	aaagccgctt	ctaaataatg	1440
cagatgtcag	tgctttgacc	ctggaataaa	aactgaaatg	acttagtgat	ttcaaaaaaa	1500
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	a			1531

<210> 33

<211> 2090

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (967)

<223> n equals a,t,g, or c

<400> 33

atagggaaag	ctggtaggcc	tgcaggtacc	ggtccggaat	tcccgggtcg	acccacgcgt	60
ccgagctgat	gcccataatt	gtattgatcc	tcgtgtcatt	attaagccag	ttgatggtct	120
ctaatoctcc	ttattcctta	tatcccagat	ctggaactgg	gcaaaactatt	aaaatgcaaa	180
cagaaaactt	gggtgttggt	tattatgtca	acaaggactt	caaaaatgaa	tataaaggaa	240
tgttattaca	aaaggtagaa	aagagtgtgg	aggaagatta	tgtgactaat	attcgaaata	300
actgctggaa	agaaagacaa	caaaaaacag	atatgcagta	tgcagcaaaa	gtataccgtg	360
atgatcgact	ccgaagaagg	cagatgcctt	gagcatggac	aactgtaaag	aattagagcg	420
gcttaccagt	ctttataaag	gaggatgaac	tgggaattttt	atttataacct	tttagcgtac	480
tcctttatctt	ttctgtaagt	aagtttggtt	tcacatcatg	ggatgaagga	aaagatttga	540
tactgaaaac	taaaactgaat	agttggttcc	tgaaatcttg	gactgtttat	gacctactgg	600
ctcctttaaa	tagtaactga	aaactaaaat	ggaatatctt	agttaacgct	tctacaagta	660
ttttcatttt	aaaagcctac	atgattccta	actaaagtgt	catgagaaag	gattatcaca	720
cctgtagcaa	ttccagttt	tagtgattct	ccattttttc	ccttgatcatg	ttaattattta	780
tggaaatgatc	attttgtgta	catacaggtt	actgcttttt	tatttaaatt	cttttagtgt	840
ttagctccat	gagacacttc	agtttaaatt	gatggaataa	atgttatatg	acacatttac	900
attttcctta	tcaaggtgtc	aaatatgtgg	actttaaaca	atgaaacttt	ttcaaaaaga	960
aaaamcnaaa	actttaactt	tgtgtaaaat	cttatagtat	tatcagctta	gaggggaattg	1020
atatttttaa	tattgccgtt	atattccaaa	atatatatgt	agataaatga	actggtgtag	1080
aatatcagtt	tgctattttg	ttttatgaat	tactatacat	atacatgcat	agaaatgaaa	1140
tgctatactg	ataaatttta	aagaaaatat	gaggaaatgg	ctataaatat	taaactaaaa	1200
gggtcttcaa	cagtaaagtg	cagttatgtc	atttaaaatt	ccaatacttt	aaaggccacc	1260
aaattttgat	gtatatgtcc	ttgaagggtc	gctaaaattt	atgaagagga	ctcacatttt	1320
cccccataga	aattttgcagt	ttcttggtga	tcatttaagc	aggatccaaa	gaagttcctt	1380
tacaaataag	taataagaaa	aatgagtact	aaaatacagc	tttgtgcctt	ttaaccctat	1440
gccaactcct	aaacatataa	gtagattaca	gtatacttat	ctgatcagag	catgayctgt	1500
ttggccacat	gcaagtgtga	gcagaaatag	agcagcacgt	agaatagtaa	cttaaagcaa	1560
gtcatccttt	aaaaattctg	agctaaaatc	tatttaccat	tgagtaattg	aattaatccc	1620
ataggaataa	gctccttgta	agtaaatcca	tgatatgaat	tagaaaaaaa	aaacagctgg	1680
aaattgaagt	ttttgatgcc	tgtatactgg	atatgaaact	atgtgatttc	tagtcttctg	1740
tgtttagcag	ttgtaatat	ttaatgattt	tgcttcatac	tcrgttaatg	gaacataaac	1800
atatctttga	tacttctttg	tgagttagag	aatgctagat	aggggtggctt	tggtctttgt	1860
ttaagttttt	tttcctgaat	gtagttaatt	tatggcatct	gttgaataaa	actgctaaaa	1920
tgacctctta	aaaatgttct	gttgatccc	ctttccagg	tgaatcaata	gaaatgcctg	1980
attgaattag	taggttaaac	taaacaacat	actgtcatag	gaaaactgga	gagcttaacc	2040
aacttgctct	tagaaatgtt	accttaaaaa	aaaaaaaaaa	gggcggccgc		2090

<210> 34

<211> 1006

<212> DNA

<213> Homo sapiens

<400> 34

gctcgtggcg	ctggaccgca	tggagtacgt	gcgcaccttc	cgcaagcgcg	aggacctgcg	60
cgccgcctg	ttttgggtgg	cgctggacct	gctggacctg	ctggacatgc	aggccagcct	120
gtgggagccg	ccgcgtccg	ggctgccgct	gtgggcccag	ggcctcacct	tcttctactg	180
ctacatgctg	ctgctggtgc	tgccgtgctg	ggcgctcagc	gaggtcagca	tgcaggcgca	240
gcacatagcg	ccgcagaaga	tgatgctgta	ccgggtgctc	agcctcgcca	ccgtcaatgt	300
gggtggccgtg	ctggcgcgcg	ccgccaacat	ggcgctgttc	cgggacagcc	gtgtctcgcc	360
catcttcgtc	ggcaaaaacg	tgggtggcgt	cgccaccaag	gcctgcacct	tcctggagta	420
ccgcgcgcag	gtgcgcgact	tcccgcgcgc	tgcgctatca	ctggagctgc	agccgccacc	480
cccgcagcgc	aactcggtgc	cgccgcgcgc	gccgctgcac	ggcccgcctg	ggcgccccc	540
catgtcctcg	cccacgcgtg	acccctgga	cacgtgacag	ggcccgcgcg	gcccccgaca	600

19

cgccccctggg	gcgcagagac	accggggttg	cttggggcgc	gcggtttgca	tgggatgggg	660
tgggggcggg	ctcccctagg	gacaggtgcc	tcgagtcccc	gtgcctgggg	tcccgcggcc	720
gcttcttcat	ctcaggaatc	tctcggaccg	cggatcctca	gcccccgctc	caccagcccg	780
ccccagsgcg	tgggtctggt	tgggaggcct	ggggcggagc	agagcagagg	tgatccggcc	840
cctgcctgct	gggcccggcg	ggttggaagg	gagggcagtg	tgggcggaga	tctgctcctt	900
cggtagggggc	ctctggtcca	gatttggggc	caaggaggcc	tctgtcattt	taaagactcg	960
tgtttacagt	tttgtaaaaa	aaaaaaaaaa	aaaaaaaaaa	ctcgag		1006

<210> 35

<211> 1787

<212> DNA

<213> Homo sapiens

<400> 35

gcagtgttgc	acttttctac	aattttggaa	aatcttggaa	atcagatcca	gggattatta	60
aaagsracaga	agagcåaaag	aaaaagacaa	tagttgaact	tgcagagaca	ggaagtctgg	120
acctcagtat	attctgcagt	acctgtttga	tacgaaaacc	ggtgaggtcc	aaacattgtg	180
gtgtgtgcaa	ccgctgtata	gcaaaatttg	atcatcattg	cccatgggtg	ggtaactgtg	240
taggtgcagg	caaccataga	tattttatgg	gctacctatt	cttcttgctt	tttatgatct	300
gctggatgat	ttatggttgt	atatcttact	ggggactcca	ctgtgagacc	acttacacca	360
aggatggatt	ttggacatac	attactcaga	ttgccacgtg	ttcaccttgg	atgttttggga	420
tgttcctgaa	cagtgttttc	cacttcatgt	gggtggctgt	attactcatg	tgtcagatgt	480
accagatatc	atgttttaggt	attactacaa	atgaaagaat	gaatgccagg	agatacaagc	540
actttaaagt	cacaacaacg	tctattgaaa	gccattcaa	ccatggatgt	gtaagaaata	600
ttatagacct	ctttgaattt	cgatgctgtg	gcctctttcg	tcctgttatc	gtggactgga	660
ccaggcagta	tacaatagaa	tatgaccaa	tatcaggatc	tgggtaccag	ctggtgtagc	720
gacatcttat	cctatgaagc	atattgctga	gtggtgcctg	aaaatttgtt	ctgtccgtgt	780
ctttctcaca	ctcgaatcca	catcctttga	acaagagcat	gctatgtgta	gggctaawgg	840
tgaattttac	agtctttttt	tcaacacttt	tattamcaa	agtaaactg	gacagaacac	900
actgcccatt	tctgggaaga	gtaaagatga	taaaaataa	ttttaatggt	tcttaatgtg	960
gaaattcaca	acataactca	cttttgggtt	ttgttctcac	agtatttttc	acaaaaaaag	1020
ggtaaactta	ttctattgac	agacatgggt	tactgatcag	aatgttccag	ttttaactaa	1080
aactaaattt	atgttatttg	gctaaatggt	atgatgcagt	ctagtacgag	tattgcatct	1140
aattccagga	gcattgtttt	aagttgattg	actagttatt	atgtacattt	cagaatgtac	1200
acataaatac	tgtgatgaaa	atcatgtgat	tgggatctac	tgtgatgttg	tcttcaargg	1260
caggagaaaa	taatgttcac	aataaaatgt	gctaacaatg	ttttgtttct	atcagctgtt	1320
gcaatgctga	tatatctcta	gttcagtgaa	ataatttgta	gtaaccttac	tctgagggtt	1380
tacggtctga	taatgaagca	cttgcagtga	tatagtaagt	catgtttttt	tgttcaaatt	1440
taaaagccct	gctaattgca	tgacacacca	catagaatgt	atactagcag	atactatcca	1500
gtgaagcata	aattagaatt	taatttgatg	ttcaaaaaca	gttccatttt	taagggttaa	1560
ggtggtattt	tcaagaaaag	gcagaacaaa	taatgcaaaa	ttctcagtaa	tagtgataca	1620
tggatatact	tcctttttaa	ttctcagctg	caaaaataatt	gtagrcaaaa	twatggcatt	1680
taactaaaaga	tggagcatga	tctgtgtaca	tagcacatgt	gaataaaaaga	aaagctgaca	1740
gtataaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaggg	cggccgc		1787

<210> 36

<211> 1201

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (29)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (48)

<223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (63)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1201)
 <223> n equals a,t,g, or c

<400> 36
 taggcttttg caaaaagctt tttaggtgnc ctatagaagg tacgcctnca ggtaccggctc 60
 cgnaattccc ggtcgacccc acgcgtccga aggaaactac ttgagraggg acccaacttt 120
 ccgctatctt ttgggttcat tccaaatagt tttgtgccat tgaaaaactt gaccttcaaa 180
 aaaatttggt tttcagaata gaacacaata ggacagtgac tgcacagttg tgaaaaagga 240
 agagaatcat taaagaaaaa gaaaaaagat ttttaagaccg ttgaaatcaa ttatcaagaa 300
 cgtcctaaaa cacctatggc tttgactttg ttattgatcc agattatatt ccttgcatg 360
 gggaaaaatc ctttcatatt tgtttgctgt aaagatgggt ttgcaagaat aagtcatgac 420
 caagacaaac tgccaataca aaagcccact gatactaatt atataatgag aaaaaaatgt 480
 atccaactag gacacatatc ttttgagtta tttggactga aagcttaaga aaacttggaa 540
 aattctatatt tgtgatctag tcaagccaca gttatcaaag gctacatttt cagtgtgaaga 600
 taaatggatg agtaaaactca aatatgtatc acgtgtgctt tgtatcttaa gatgtgtttc 660
 caagagcatc tgaaattttg tttgtacatg tatcttgatc atttataaag ccactgtgat 720
 ctataaatca agaaaaatcca ttgtcataac cattttttaa agtcaaaaat taagacatcc 780
 ttaattaaaa agtttcaaact ctagacacta aatgtgtgtg aatgtacaaa gaaaacaaac 840
 cattgcttat gctgttatat actagagaaa ttttggtttg cttgctgttt taacttgaca 900
 gatgaaggac tttagtgtgaa cttcatattg taagaactgt taataaaaagt tgtcaagtaa 960
 aaagcgctat atctaaaaag actttatgaa cagttattct atcaactttt aaagggtttta 1020
 aacctgcccc gaaattacct tggatatctga agtttccctc tgtctcctcc tctaattaag 1080
 cttgttatatt gtcatgcacc agcattggag ataataaaaat ttcttgttct gtgtaaaaaa 1140
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1200
 n 1201

<210> 37
 <211> 1896
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (444)
 <223> n equals a,t,g, or c

<400> 37
 ctgcaggaat tcggcacgag cggaaccggg gccggctgct gtgcatgctg gcgctgacct 60
 tcatgttcat ggtgctggag gtggtggtga gccgggtgac ctcgtcgctg gcgatgctct 120
 ccgactcctt ccacatgctg tcggacgtgc tggcgctggt ggtggcgctg gtggccgagc 180
 gcttcgcccg gcggacccac gccaccaga agaacacgtt cggctggatc cgagccgagg 240
 taatgggggc tctggtgaac gccatcttcc tgactggcct ctgtttcgcc atcctgctgg 300
 aggccatcga gcgcttcacg gagccgcacg agatgcagca gccgctggtg gtccttgggg 360
 tcggcgctggc cgggctgctg gtcaacgtgc tggggtctg cctcttccac catcacagcg 420
 gcttcagcca ggactccggc cacngccact cgcacggggg tcacggccac ggccacggcc 480
 tccccaaagg gcctcgcggt aagagcacc gccccgggag cagcgacatc aacgtggccc 540
 cgggcgagca ggtcccgac caggaggaga ccaacaccct ggtggccaat accagcaact 600
 ccaacgggct gaaattggac cccgcagacc cagaaaaccc cagaagtggg gatacagtg 660
 aagtacaagt gaatggaaat cttgtcagag aacctgacca tatggaactg gaagaagata 720
 gggctggaca acttaacatg cgtggagttt ttctgcatgt ccttgagat gccttgggtt 780
 cagtgattgt agtagtaaat gccttagtct tttacttttc ttggaaagggt tgttctgaag 840

21

gggatttttg	tgtgaatcca	tgtttcctg	acccctgcaa	gccatttgta	gaaataatta	900
atagtactca	tgcatcagtt	tatgaggctg	gtccttgctg	ggtgctatat	ttagatccaa	960
ctctttgtgt	tgtaatgggt	tgtatacttc	tttacacaac	ctayccatta	cttaaggaat	1020
ctgctcttat	tcttctacaa	actgttccta	aacaaattga	tatcagaaat	ttgataaaaag	1080
aacttcgaaa	tggtgaagga	gttgaggaag	ttcatgaatt	acatgtttgg	caacttgctg	1140
gaagcagaat	cattgccact	gctcacataa	aatgtgaaga	tccaacatca	tacatggagg	1200
tggtctaaamc	cattaaagac	gtttttcata	atcacggaat	tcacgctact	accattcagc	1260
ctgaatttgc	tagtgtaggc	tctaaatcaa	gtgtagttcc	gtgtgaactt	gcctgcagaa	1320
cccagtgctg	tttgaagcaa	tggttgggga	cactaccaca	agccccttct	ggaaaggatg	1380
cagaaaagac	cccagcagtt	agcatttctt	gtttagaact	tagtaacaat	ctagagaaga	1440
agcccaggag	gactaaagct	gaaaacatcc	ctgctgttgt	gataagagatt	aaaaacatgc	1500
ccaaacaaac	aacctgaatc	atctttgtga	gtcttgaaaa	agatgtgata	tttgactttt	1560
gctttaaact	gcaagaggaa	aaagactcca	ctgaaattct	aagtttgcca	agtagtgtaa	1620
ttgaagtctt	tgctctgtca	cacagtttaa	ttctattttt	gtaagaacat	aatgggactg	1680
cataacagag	ttctatatta	caattttgtg	attattagta	cagagtacag	ctatgctgtg	1740
actgttttgg	aaagccagtt	ttaacactat	gttacctttt	tgtttaaagt	aagttaaacc	1800
ttatataaca	taatgacatt	tgatttctgg	atttttccca	tgataaaaat	tagggggata	1860
aataaaattg	ttactggaat	ttctctgcaa	aaaaaa			1896

<210> 38

<211> 1152

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1145)

<223> n equals a,t,g, or c

<400> 38

agttccagga	taaaaacaga	ccgtgtctca	gtaactggcc	agaggatacg	gatgtcctct	60
acatcgtgtc	tcagttcttt	gtagaagagt	ggcggaatt	tgtagaaaag	cctacaagat	120
gcagccctgt	gtcatcagtt	gggaacagt	ctcttttgtg	tccccacggg	ggcctcatgt	180
ttacatttgc	ttccatgacc	aaagaagatt	ctaaacttat	agctctcata	tgcccagtg	240
agtggcaa	gatacaaaag	ctctttgttg	tggtatcatgt	aattaaaatc	acgagaattg	300
aagtgggaga	tgtaaaccct	tcagaaacac	agtatatattc	tgagcccaaa	ctctgtccag	360
aatgcagaga	aggcttattg	tgtcagcagc	agagggacct	gcgtgaatac	actcaagcca	420
ccatctatgt	ccataaagtt	gtggataata	aaaagggtgat	gaaggattcg	gctccggaac	480
tgaatgtgag	tagttctgaa	acagaggagg	acaaggaaga	agctaaacca	gatggagaaa	540
aagatccaga	ttttaatcaa	agcmtggtg	gaacaaagcg	gcaaaagata	tcccatcaaa	600
attatatagc	ctatcaaaag	caagttattc	gccgaagtat	gcgacataga	aaagttcgtg	660
gtgagaaagc	acttctcgtt	tctgtcta	agacgttaaa	agaattgaaa	attcagatca	720
tgcatgcatt	ttcagttgct	ccttttgacc	agaatttgct	aattgatgga	aagattttta	780
gtgatgactg	tgccacccta	ggcacccttg	gcgtcattcc	tgaatctgtc	attttattga	840
aggctgatga	accaattgca	gattatgctg	caatggatga	tgatcatgca	gtttgtatgc	900
cagaagaagg	gtttaaaggt	actggtcttc	ttggacatta	atctttgaat	acttgctgac	960
tgctaagaaa	tgaccagagg	ggaagaggag	tttgacatgt	tagggcatta	aagcaaaggt	1020
ggatttaaga	attaaacat	tacatgcccc	ttccaaaagg	cagaaatcca	ttcaaactgtg	1080
actgtcccaa	atgccttatg	tcaaataaag	cagattgcac	tgatggaaaa	aaaaaaaaaa	1140
aaaanactcg	ag					1152

<210> 39

<211> 1017

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (822)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (994)

<223> n equals a,t,g, or c

<400> 39

gaacaaagt	cagtactga	gagggctgag	cggaggetgc	tgaaggggag	aaaggagtga	60
ggagctgctg	ggcagagagg	gactgtccgg	ctcccagatg	ctgggcctcc	tggggagcac	120
agccctcgtg	ggatggatca	caggtgctgc	tgtggcggtc	ctgctgctgc	tgctgctgct	180
ggccacctgc	cttttccacg	gacggcagga	ctgtgacgtg	gagaggaacc	gtacagctgc	240
agggggaaac	cgagtccgcc	gggcccagcc	ttggcccttc	cggcggcggg	gccacctggg	300
aatctttcac	catcacctgc	atcctggcca	cgatctcat	gtgccgaatg	tgggcctcca	360
ccaccaccac	cacccccgcc	acamccctca	ccaccwccac	caccaccacc	acccccaccg	420
ccaccatccc	cgccacgctc	gctgargctg	ctgtcgcggg	tgcctgtgga	cagcagctgc	480
ccctgccttc	ccatctgttc	ccaggacaag	tggaccccat	gtttccatgt	ggaaggatgc	540
atctctgggg	tgaacgargg	gaacaataga	ctggggcttg	ctccagctgc	atttgcatgg	600
catgccccag	tgtactatgg	cagcagagaa	tggaggaaca	ctgggtctgc	agtgtgaag	660
ggtttgggga	gtggagagca	aggggtgctc	ttcggggctg	gacagcccg	cttgtgacag	720
tgactcccag	tgagccccag	aaatgacaag	cgtgtcttgg	cagagccagc	acacaagtgg	780
atgtgaagtg	cccgctctga	cctcctcatc	aggctgctgc	angcctctgg	cgggcagggc	840
actgggagag	gccctgagaa	tgtccttttg	gtttggagaa	ggcagtgtga	ggctgcacag	900
tcaattcatc	ggtgccttag	tccaagaaaa	taaaaccac	taagaaaaaa	aaaaaaaaaa	960
aatgaccctc	gagggggggc	ccggtaccca	attngcccta	tgaagaggcg	aacagga	1017

<210> 40

<211> 1777

<212> DNA

<213> Homo sapiens

<400> 40

ggcacgaggt	ccccgacgcg	ccccgccccaa	cccctacgat	gaagagggcg	tccgctggag	60
ggagccggct	gctggcatgg	gtgctgtggc	tgcaggcctg	gcaggtggca	gccccatgcc	120
caggtgcctg	cgtatgctac	aatgagccca	aggtgacgac	aagctgcccc	cagcagggcc	180
tgcaggctgt	gcccgtgggc	atccctgctg	ccagccagcg	catcttccctg	cacggcaacc	240
gcatctcgca	tgtgccagct	gccagcttcc	gtgcctgccc	caacctcacc	atcctgtggc	300
tgcactcgaa	tgtgctggcc	cgaattgatg	cggctgcctt	cactggcctg	gccctcctgg	360
agcagctgga	cctcagcgat	aatgcacagc	tccggctctgt	ggaccctgcc	acattccacg	420
gcctggggccg	cctacacacg	gtgcacctgg	accgctgcgg	cctgcaggag	ctggggcccg	480
ggctgttccg	cggcctggct	gccctgcagt	acctctacct	gcaggacaac	gcgctgcagg	540
cactgcctga	tgacaccttc	cgcgacctgg	gcaacctcac	acacctcttc	ctgcacggca	600
accgcatctc	cagcgtgccc	gagcgcgcct	tccgtgggct	gcacagcctc	gaccgtctcc	660
tactgcacca	gaaccgcgtg	gcccattgtg	accgcgatgc	cttccgtgac	cttggccgcc	720
tcattgacact	ctatctgttt	gccaacaatc	tatcagcgct	gcccactgag	gccctggccc	780
ccctgcgtgc	cctgcaatac	ctgaggetca	acgacaaccc	ctgggtgtgt	gactgccggg	840
cacgcccact	ctgggcctgg	ctgcagaagt	tccgcggctc	ctcctccgag	gtgccctgca	900
gcctcccgcga	acgcctggct	ggccgtgacc	tcaaaccgct	agctgccaat	gacctgcagg	960
gctgcgtgtg	ggccaccggc	ccttaccatc	ccatctggac	cggcagggcc	accgatgagg	1020
agccgctggg	gcttcccaag	tgtgtccagc	cagatgccgc	tgacaaggcc	tcagtactgg	1080
agcctggaag	accagcttcg	gcaggcaatg	cgctgaaggg	accgcgtgcc	ggccggggac	1140
agggccggcg	ggaaacgggt	tttggcccaa	gggaacatta	atgacttacc	cttttgggac	1200
tctgcctggt	tttgggtgagc	ccccggttac	ttgcagtgcg	gcccagaggga	tccgagccac	1260
caggttcccc	acttcggggc	cttcggccga	ggccaggctg	ttcacgcaag	aaccgcaccc	1320
gcagccatgc	cgtctggggc	aggcaggcag	cgggggtggc	gggactggtg	actcagaagg	1380
ctcaggtgcc	ctacccagcc	tcacctgcag	cctcaccccc	ctgggcctgg	cgctgggtgt	1440
gtggacagtg	cttggggcct	gctgaccccc	agcggacaca	agagcgtgct	cagcagccag	1500
gtgtgtgtac	atacgggggc	tctctccacg	ccgccaagcc	agccggggcg	ccgaccctgt	1560
gggcaggcca	ggccaggctc	tccctgatgg	acgcctgccc	cccgccaccc	ccatctccac	1620

23

cccatcatgt	ttacagggtt	cggcggcagc	gtttgttcca	gaacgccgcc	tcccacccag	1680
atcgcggtat	atagagatat	gcattttatt	ttacttgtgt	aaaaatatcg	gacgacgtgg	1740
aataaagagc	tcttttctta	aaaaaaaaaa	aaaaaaa			1777

<210> 41

<211> 1003

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (990)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1002)

<223> n equals a,t,g, or c

<400> 41

aattcggcac	gagttcctct	cctcctgttt	tgctacatto	tcctcagtgg	caaaaagttt	60
cactctacct	ctgacagcat	gtatattgca	ccagtagcta	acaaaaactg	gtctagtcaa	120
accaaagtgg	cacaaaagaa	ccaggatacc	aaaagttaag	ctcatacagc	tgcaaaccat	180
atcacttctt	ggtaacaatg	cagacctcat	aaacctaaag	aagagaaaga	aaagaaaact	240
tttgttactt	tccttttttg	cttgctcactt	atatacaggc	tatgtgagaa	tataatttgt	300
aggtataaca	cattaagaaa	aagttatctt	cattggatag	aattgaatgg	tggtcgctga	360
taggaatagg	gcgtcctcta	gctcttatct	ctgtctctta	ctcttttctc	tttctctttt	420
tctctgtcat	gagactgtgt	gtgacagggc	cacctgtctt	tttttttttc	ttaaattttt	480
ttttcttttt	atgtgtaggt	gcatgtcttg	gggatttaaa	aatttcaagg	ctgggtttact	540
tatgcaaagc	atgcctacgt	ctggaatact	tagggaaaga	aagcgactcc	atgttgctcg	600
aattcctcaa	gggacagaaa	aaaaattgga	gactgttgaa	atgcagattt	gaagtaattt	660
ttttaaaata	ttattttggg	ttctgcgaca	ttgtgaaaaa	ttaaagtgtg	tgtgcaatac	720
ttaattcaga	catgtaccac	aagttaatgg	tagactaaca	ctgggggggtg	gggtctaggc	780
atcatgcttt	tgctcagcata	ctcttgagct	tttaagtcta	ctatgtctga	actgtggttt	840
cttgttttatc	cttttttctt	tagttggact	gtaatgtatg	gtctgtcaac	ctgtgaatct	900
ttaaagtatg	attcaggtat	tgttgtattc	tttactgtgt	aataaaaaag	ttgaaaaaaa	960
aaaaaaaaaa	acccaagggg	gggcccggtn	cctttccccc	tnt		1003

<210> 42

<211> 1201

<212> DNA

<213> Homo sapiens

<400> 42

ccttcactga	gtttttctta	tctccttttg	cagtgggttaa	ctattgtcac	aataatgctg	60
aatgaaaaaa	cacaccaa	gtcagtggct	taaaatgacc	atcctttttc	tcacagttat	120
gaagattggc	tatggcacct	ctgcttcttg	ctataggcct	gaggttctag	ggcttcttat	180
gcctcatcct	ctttaagcca	aagggatagc	cagagcatct	tgatggcaga	agtgaataa	240
gatgagcccc	actgctcggg	tacattttca	gccctgggtt	gtgtcatgtc	tactgatata	300
tcattgggtca	aaggaagtca	gagggccaag	atgaagaggc	agggaaatat	gcactgccca	360
cagtgaagcc	atgacaagag	tgaggatgca	ggaaggcatg	aagaattggg	gccaacagtt	420
caatctacca	taccttctct	cacctggaat	tccagatgct	tgagctacga	aacttagatg	480
caaagaaagt	taaagctaga	aggaacctca	ggcccagttg	ctcatttttg	agattccaaa	540
tgtgaatttc	agagagctga	gataaacttg	ccaaggccat	atagaggctg	tgactaaatc	600
tggaacttaa	tccagactat	caatcttagg	ccagtgttct	tttttcaata	tagtccttgg	660
cataatgcta	tgcttattag	gtagataaaa	gggcttatgt	caagaaattt	ggagcagagt	720
ctgattactt	gagcatgaac	atacccagac	aaggtatgtt	ctggagtcac	attctagcct	780
ctgagctcat	tttttcatgc	gagttcatat	aaaatcctcg	aaagttaga	aactaggttt	840
tagtagtaac	ggagctagaa	tcattcttcg	gcttattcct	gctagttgtt	ccatatttct	900

24

agatttcac	ttgaattttg	aaaactgatt	taagaatata	tttagtatta	ttattagtaa	960
gggaatacgc	aatccagttt	caatttttatt	cagaagtagg	tcacctaatt	ctagaaaatg	1020
gttattagtc	tagtgctgct	tagcaaggta	cttaaaaagaa	aatctgcaca	tatccttggtg	1080
ctgcccttct	taaaaacaga	aaacaaaaag	tgtaagatca	tcattgcttc	ccacatagga	1140
aaaataaaat	gtcttcagac	ttgatgtgaa	aaaaaaaaaa	aaaaaactcg	aggggggggc	1200
c						1201

<210> 43

<211> 1176

<212> DNA

<213> Homo sapiens

<400> 43

tttgattggt	ttgtaatgct	caagtttctg	tcattttcaa	atatgttggg	cttggtcttc	60
atggcaaacc	tagaatccta	gactgccgct	ccccagagga	ggttctttaa	gactgctcag	120
ctctcctgcc	aatagcaaca	atgcaaaagc	ttacccttcc	tcccgtttcc	ccagcccat	180
cttcattgtc	tgtggatggt	gcttcaccca	catttataat	ttactcctgt	ctctctgcta	240
tgggttgggt	gttgaaagag	tgaagttctt	tacctttagt	attttaaaaa	aagaaacaat	300
gttgctcaat	tatttattct	aaatatgttg	ttgggctttg	ctttatttac	cctggttgca	360
tgtcttggtg	atgtcttttc	agtccttatg	gcccacgcac	ttccttatcc	atgaccaga	420
gaggccccag	tgatattttg	acttttcaaa	tgtggtgaat	aagtgagagt	tgtttggtga	480
gttaactgtg	attttaaata	ttctgattgt	tgtgaggcac	ttttctaggt	gtttgatttc	540
ttgatctggt	ttcttctatg	ccaattgatt	aaacagtgtc	ttccacagtt	tgctaagggtg	600
atgatggtgc	ctggttggtg	gttttggttg	gttgaataaa	gccctgatgc	tgggagttca	660
ttggtgtagg	tgggtcaagt	ggtggagtcc	tgtaagcatt	cctgtgtatt	cttctctgaa	720
taataatgct	gcatattaag	cctagcatcc	taccatatt	accacataaa	ctgttaggtc	780
tgtcctggct	tcaattcatg	ttggcggtca	ccttaaat	taaaaaataaa	gtttatgttt	840
atttgagggg	caagtgacta	tgttgtaaga	atgatatttt	tctgaccagt	agttttattt	900
tatttttact	ttttatttgt	tccaagatgg	agtcttgctc	tgtcaccacag	gctggagtgt	960
agtgcacaaa	tctcggctga	ctgcaacctc	cacctcccg	gctcaagaaa	ttctcctacc	1020
tcagctactc	gggaggctga	ggcaggagaa	tcgcttgaa	ccgggaggcg	gagggtgcag	1080
tgagtcgagg	tcgcaccatt	gcactccagc	ctgggcaaca	agagcaagat	ccgtctcaaa	1140
aaaaaaaaaa	aaaaaaaaaa	ctcgaggggg	ggcccg			1176

<210> 44

<211> 569

<212> DNA

<213> Homo sapiens

<400> 44

cccgggtcga	cccacgcgtc	cggcaggcag	caggaagga	agcagaggct	tcctaaggct	60
gttttcttag	cgttgagaa	gcccgcgctt	tctacatgct	cccaagtgtc	gtcatgagca	120
cgttcctgac	aagtcagggtg	ttcagattgc	agtccctggc	caacgtcagg	attcttacag	180
gttgaatggt	aagctcaccg	atcttggcct	caggtoctgc	ctggcttgcc	tgctcatttt	240
cacacgtgca	gtggtgggtc	tgtctcatag	cacagggtgca	gtttagtgc	gccacagtgt	300
ccccaggcag	ggcggggact	ggagcagccc	ccagtgtgcc	agcagtgtgg	gcagcggagg	360
ctaggggccc	atctgtccct	agcaccctcc	agggcagtc	cgttttgag	cgggatttgg	420
caaaccccc	tccaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	480
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	540
aaaaaaaaaa	aaaaaaaaag	ggggggccc				569

<210> 45

<211> 986

<212> DNA

<213> Homo sapiens

<400> 45

gcactggcct	cttcactggt	ggccgagaa	cagggctttg	tggcagcact	gatgggtgcag	60
gaggcaccgg	ccctggtacg	gctgagcctg	gggtcccatc	gggtcaaggg	cccactccca	120

25

gtgttgaagc	tccagccgga	gggctggagc	ccatctactc	tctggagctg	cgtttccgtg	180
tggaaggaca	gctgtatgca	cccctggagg	ctgtccatgt	gccctgcctg	tgtcctggcc	240
gccctgccc	ccctctgtc	ctgcctctgc	agccccgatg	cccggccccc	gcacggctgg	300
atgtccatgc	cctttacacc	acatccactg	gtctcacgtg	ccatgccac	ttgccacccc	360
tgttcgtgaa	ctttgccgac	ctctttctgc	ctttcccga	gcctccagag	ggggccgggc	420
tggtcttctt	tgaggagctc	tggtattcct	gcctgccaga	gggtgctgag	agtcgtgtgt	480
ggtgtccact	tggtccacag	ggcctggagg	gcttggtgtc	ccgccacctg	gagccttttg	540
tggtggtggc	ccagcctcct	accagctact	gtgtagcaat	ccacctgcc	ccggactcaa	600
agctgctgct	gcggctggag	gcggccctgg	cagatggagt	gcctgtggcc	tgccgaccga	660
tgactggggc	gtgctgcccc	tgccggggga	ctacctccgt	gggctggcgg	ctgctgtctg	720
agccccggga	gaccaggtgg	gggcaggact	gtggcccttg	tgggggccaa	ggcacactcc	780
tgtagctctg	tcgccaaaac	cctgcattcc	gcagtgcctt	cgctggcttg	ttttcttttg	840
ggccccgggt	gggagcaggc	tcctgggggt	gagggctctg	ctgagctctg	ttttgctgct	900
ctagcaagat	ccctgagacg	gggtaagtta	taataaacag	aaatgtattg	gtcagaaaa	960
aaaaaaaaa	aaaaaagggc	ggccgc				986

<210> 46

<211> 1540

<212> DNA

<213> Homo sapiens

<400> 46

ggcacgaggg	aactagtata	ttcacgtct	atgaggccgc	ctcacaggaa	ggctgggtgt	60
tcctcatgta	cagagcaatt	gacagcttcc	cccgttggcg	ttcctacttc	tatttcattca	120
ctctcatttt	cttcctcgcc	tggtttgtga	agaacgtgtt	tattgctgtt	atcattgaaa	180
catttgcaga	aatcagagta	cagtttcaac	aaatgtgggg	atcgagaagc	agcactacct	240
caacagccac	caccagatg	tttcatgaag	atgctgctgg	aggttggcag	ctggtagctg	300
tggtgatgtca	acaagcccca	gggacgcgcc	ccagcctgcc	tccaggtgca	gtacaatgac	360
atttttaaaa	atcgcccagc	aaaggctctt	gaattttatt	tcattccaaga	aaatccacag	420
ctctttaagc	tctagatttg	tccaaattta	aaatcctgaa	gttagagatg	gtatttctact	480
ccttcctcta	ttcccaggac	ctagcttttt	ttttttaaca	tacacaatag	ggatttgata	540
agtttctgat	ggctgcaggc	atgtaagagc	atttcagtgg	tattgaatca	atgaagaatt	600
ttgttgacat	gtgaaatctt	ataaaaaatat	tctttaccga	aggactgagt	tatgtggcag	660
tgggcaaat	cattgtttca	tacctcccct	agtaactggg	aaaaatatgt	taatacatag	720
tctctctgtt	tttctgcatt	tggaagcttt	cagaggaaca	taatgtagag	gtgtttcttt	780
agcaaatgtc	actgatatga	aacataagga	ttgcagggtg	ggcctgagag	tcctcatgag	840
atagattctc	acagtgatta	gaagatggag	tctcacgtcc	ctgcctgtga	actttctgga	900
aaaaccatct	tctccaagct	gccattgaca	acaatatgga	taacaataat	aacaataagg	960
cccaataaac	tcctttatct	cttcttcagg	gggcatact	gacatcttct	cttccttggt	1020
ttcccctcct	tgccccctaa	atatccagta	actcattcaa	aataatgtca	ccttaccag	1080
agcagcccc	ctaactttcc	ataatatttt	cactttcatt	ttcccctcaa	gcagcccact	1140
cgtaggaccg	tagaattgat	tcttccacct	ggagaatttt	attttcttta	gccttttttg	1200
ttttcagtga	caaatcctct	tctcgcaagg	ggtgggtttcc	atagttgttt	atatcctgcc	1260
ctcataattt	ggagaagtgt	tcacatctgc	cgtgggatga	gactgtatct	cttttctttc	1320
ttttgggtct	ttctccagat	agggacttct	tatgcaactc	aaggatgggt	acatgaaaa	1380
taaaattgta	ctctgagcca	ttactggtgg	gctatgttta	tatggccatt	ttaccataga	1440
gttattttact	tctttttgtt	tctattttgta	ttgaggtgtg	attaacaaat	aaaattgtaa	1500
atacttaagg	taaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa			1540

<210> 47

<211> 792

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (759)

<223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (760)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (774)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (779)
 <223> n equals a,t,g, or c

<400> 47
 actttccagc taaaaaccaa caagtgtctg aggacacagt ttaaactcca agatgatagg 60
 gtcctccctc acctgggctc ccacctaccc tcatgacctc cttttgtgaa atgctgaagg 120
 gctctgcagc tggttgtctg gtactgctgg cctttgcttt ctatttagca tgctccttct 180
 cccacaaaac aaaatcacat tctcactatg ccctgttcat tcttcaggac tatcttcttg 240
 gaaactttta ctacataccc ctctccccct aatctgagtg tctgctttgc tcaggtagca 300
 tgtgttcaact ggataaatcc ttgattcctg gcactgaggc aggggtttctg tccccaggaa 360
 gcagaggcat actattctgt gaaggattga ctgagtttct cctaatacca agcagtatct 420
 gagggaacag atgtctagct taaaatcctc cctagcactt gtcatagcag tgctacgtat 480
 tgctgtgaa ggaagttaa taactgtga aagggttcgat tagctttatt tcctcaggat 540
 ttgtttgact ttacaaattg atttgggtta ttycaacttt taggtctagt cttaagtata 600
 actggtacat attccttcaa gcagccatta cacctctcat aaatttatta tacacctgca 660
 tttttataac tattatgctt ttttaattgtt ggccaccatt tttagtgtct ctgaattggt 720
 atggttctca agcagcagtt gtcaccttg ttttgaatnn atgctgtgac ggangcctnc 780
 aggggaattc cc 792

<210> 48
 <211> 1497
 <212> DNA
 <213> Homo sapiens

<400> 48
 gtgtaaaacc agtttcgggt tagccatgtc cggtgcacac atgcatgcat gtgtgttctt 60
 gtgcgcgtgt gttttctgtt tagcagagaa tgcgctagag agcgtgatca tcctgtgcta 120
 ttcatataat aaagatgaag tgagagaaca ttagaggaac caaggccatg tgatggtaca 180
 cgtctgacgt tttttccttt cggttacatg tccgtatctc ctctttcccc tttttccctt 240
 ttgtcttcat ttggttcccc tccctatagg gagtttagga caagaagagg ctaaagtttc 300
 actgatgagc ctttctgagg gttctccatt aaatccaagg acagaaaatg tacagtccctc 360
 ttattagcat aacgaagcca tcagcattgc atcaagcggg tcctcgtacc cttttccttg 420
 taatgggtgtt tgggtgaggg tcctgaggaa gagctgccag cccctacctg atggatcaaa 480
 atccccttgg caccaaagag tgactgatag tgtaaccat cacaggagac atgtatgtat 540
 gtgtgtggaa ctctgacgtg ttttttaaac tttgcaatag cccaaagtta atttgtttct 600
 ttgccatttg ttttcgaatg ggtgtggcat tgcttataaa atgttgaatg taagttgcct 660
 aggacggcgc ctggcatatg gtggaaactg aataaaggct gctaggtagt gttagctaga 720
 tggactagaa aacagyacag atgcagatgc tttcagatgt tctctctgcc acagagagac 780
 ctttctgtgt gctttgttca aagttgacag tgtgagtaaa ctaccatcaa caaggsqgta 840
 cttttgggta attttttcaa tgkttatccc agttccttca tccgctttta catagcctct 900
 tgkactgcaa gctacactca gttttgaaga tgggtgggta gcgttagagt ggtgttctgt 960
 ggccagtggg ggaggaagct tgctcacttt aatgcagaat gtctaagcat cctgcgctaa 1020
 ccattggcag gaaggttatt tcagttagac gctggttctt tctcaccctt gcacccttcc 1080
 tggaatcacc actggttgca gaagcctata tagrgtggct catttgagc aagtattggt 1140
 aaggtgggta ttagaagact cgcaacttag aaaaggaagt aaacatgtta atactagctt 1200
 tcataaatcc ccctctctaa aaaacacccc tttctaaaaa attcacatta ctaaggcatc 1260
 ccctacacag aatgttttagt agggaggtat taaaattaat agcaattctg agtaagttcc 1320

27

ttctcatgta	gttaatgcag	catgatgaaa	gaaataaaaag	tctccttatt	tttttacttt	1380
ctgtgttgct	tcttaaataat	actccttcaa	gtcagggttag	cttaccttgg	gtgtttgtat	1440
tttgttggt	tatgtttgca	tgtgtgaatt	aaaaaaaaa	aaaaaaaaa	actcgag	1497

<210> 49

<211> 1340

<212> DNA

<213> Homo sapiens

<400> 49

ggcacgagaa	agaaaggcga	gagaaaaatc	aaggcaccaa	atttagattg	gagggtctcag	60
aggagcagtg	ttttccctcc	ttcgtaacag	ttgaacaact	tccagatgta	gctagctgca	120
ccccctgtaa	agatgcaggc	tctttacaat	gaagacacat	cttctgatgt	tccttctctc	180
ctgtatggcc	agatgcacag	gaatagtgcc	caaaagacct	cagcctgctt	tccttttaag	240
gggaaggaga	agaaaaaact	cctttttatt	tttactttct	ttcagcattg	aatttttgtt	300
gtgtgtatgg	tgacttctgt	ttttgggaaa	cggagaagc	cagcagcatg	ctgaattgtc	360
ctgacaggct	tccgctggct	cttgccgagg	ttagcagtgc	tttttttgta	tttaaaccat	420
ctcccgggca	gtgtaaaaag	tttgacagtg	cggacattct	gtctgactgg	tctcggcagt	480
gctctataac	cctgttgtgt	ttcttgataa	aacacagccc	caccctttaa	taaagcaaag	540
attgctatga	aaccagagag	tctattcatt	actgtggagt	aactagagca	gtctgtagt	600
actagacata	cggcaattag	gaagtcattg	agttgggatt	tttgtcttaa	ttttggctgc	660
tcaaagtgcc	ccctgtagga	tattcttttt	tcgggaattg	tttccaaact	tgctgtctt	720
tatctatggt	gaaactcaag	ccgcttttta	aggcaagcct	gcaaacccaa	gtatcaacat	780
gggtctctga	aggcacagg	agcagattca	cagttctgac	cagtgttagg	gtccccacga	840
gggccaccca	tttgaactca	aggttggcag	actctggccc	cagcacttgc	cgtgggttca	900
ggatggccag	cgggtgacaca	gggctatgga	accctgggtc	ttcatctctt	cccatatcct	960
ttgtttcacc	ttctttttgc	ccatatttta	ttgtgcttca	gatagaaatt	ttatttataa	1020
gataaaaagt	agctctgagg	ctgggcacgg	tggtcatgac	ctgtgggtccc	agcacttttg	1080
gaggccgag	tgggtgggtc	acgagctcag	cagatcaaga	ccatcctggc	caatatggtg	1140
aaacctgtc	tctgtctaaa	atacaaaaat	tggttggcgg	tggttggcgg	tgctgtagt	1200
cccagctact	cgggaggctg	aggcgggaga	atcgattgga	cccaggaggc	ggaggttgca	1260
gtgagcctag	atggcaccac	tgcgctccag	cctgggtgac	agaggggagac	tgcttcaaaa	1320
aaaaaaaaa	aaaaaaaaa					1340

<210> 50

<211> 1539

<212> DNA

<213> Homo sapiens

<400> 50

cgatggcccc	gcgccgctc	tagaaagtcc	cgttttttt	ttttttttt	ttttttttt	60
tttttagagta	cgttctgcat	tttatttytg	caggcaacac	tttgctcacc	agcaagaaca	120
cagcccragg	aagggaacca	ataacctttc	aaaacscaa	ctgctkctg	cgggtgagggc	180
ccagggtcct	ccacggagag	gacaggcatc	ttcctttccc	accaggaagg	agtcagccc	240
gagcctctgc	tatgtgcaag	gcggtgtgca	agcaccggct	gcrgetyttt	gctgtctctt	300
ctttctcttt	ggggctgggc	tggtgtgtgc	ttctggtgct	gatgctttgg	cctgtgaggc	360
tgagcttggc	acctcgaccc	gttcaattac	agcaacgaag	aagccactgc	tgagtgtggt	420
ctcaggggag	gcccggaggc	agtgtctggc	accggggaac	gtgctcaggc	ctcgggtggg	480
ccaggcaggc	agggcgagg	ctagcctgaa	ggcgcccggg	ttctgctgca	gcgcactctc	540
caccacgtct	tcattctcct	cctggcagag	ggagcacgtg	gagtagacga	gccgctgcag	600
ggaagggaaa	gtgagcgcgt	ggcacagggc	tcgctgctgg	aaccctgcca	gggcatgcag	660
acgcaccggg	ctaggtgtgc	ctgccccggg	ctcctccagc	tgtctgctcg	gcatacccgga	720
gccactgcag	gaaggatcca	gcaggayrta	gtggacctca	ygrtagcgyg	gatcyraggg	780
ggagaccgcc	aggaagtcct	cctcagccag	yticacagcar	gagacgccag	cccrggccag	840
cagcgtggcc	atggatgcca	gccgcttggc	atccaggctca	aaggcaaaaga	tcttcccttg	900
gttcttcaga	agagcagcca	agtgactggt	cttattgcct	ggggcggcac	aggcatcgat	960
gacatgggag	cctggcgggg	ggtccagcag	catggctggg	agacagctgg	ccctgtcctg	1020
cagaatgagg	tgtccggccc	ggtacagtgg	gtgttcatgc	agatctgtct	ggggcggaag	1080
caccagcagc	tccggcatca	aggggtccag	gagaaaatgc	ttccccttga	gggctcgtaa	1140

28

gtcatcgagg	ctggaagccc	gaccttgata	ggagaaacct	tgtctcttga	aataatcaac	1200
tacatcatcg	gagcaggtct	tgagagtgtt	cacacgcaca	aatcgaggca	gctgggaggc	1260
tggaccaggc	ctggatccca	cttccaacag	gtcctcattc	cggctcacac	cccgatgaac	1320
cttgagccga	gccaaactcag	ccttgagcct	cgcctgggtgc	cggcccaaca	gagccttcca	1380
tcggccccca	ccccctcgaa	agcccttttc	caacaacaac	tcatacacta	gcaccttggc	1440
caggtgcggc	cgctctagag	gatccctcga	ggggcccaag	cttacgcgtg	catgcgacgt	1500
catagctctc	tccttagagt	gagtcgaatg	aggttcata			1539

<210> 51

<211> 1423

<212> DNA

<213> Homo sapiens

<400> 51

ggcacgagct	tgaacatata	taatgaagaa	atacagtggc	tctttattaa	aaataatagt	60
tggataatat	aaactgaact	atztatgcat	ttttatatac	ttataaatcc	ttccaaatag	120
ttttaattct	atccttttac	atataaataa	cttaataagt	gtgctggaaa	aacacagatg	180
ttcacagcac	cactgttttt	tttttttttt	tttgagataa	taaattccat	gagaaatctg	240
ggtttgaata	tttgtttact	ttgtctccta	attgaacacc	actccaggcc	ttctgtctgt	300
ctccccctta	cccccaaaat	actcacaaaa	aaattttaag	acaacaagta	accatatata	360
gggtgttgaa	tgattttctc	atttttatct	aatttcattt	cataagtccc	gagtaattta	420
cctaccatag	gctactatac	tgataatata	aatgaaaccg	aacatttttt	gctactaact	480
ctccccaatt	taatgtgttt	tcgaaataaa	aattttaaatt	tttttccttt	taattaaaaa	540
gtcatctttg	aagtccttat	tggtgtgaca	ttttacatgt	ttgctgggtac	tattattttg	600
tcagtcagtt	aaagctggca	tgtacagctc	ttggctttta	tgaaaagcac	attgacataa	660
tgtagtaaaa	ttccaaaccc	cggcacagaa	tgtgagttaa	aattaagtct	tgctgggtta	720
gtgtacaata	aactatacct	acagactttt	ttttaataga	aagaagacaa	agctgctggg	780
ataggatttg	tcctttgaag	aaaaaatgag	ggaacaaac	acaaaaaccc	aatgcagtgt	840
ataaataaca	ttttgttcaa	ctacctctta	atgtggaatt	atctacttta	atagtttcct	900
gacagtaatg	ttaaatagta	actgccaaat	ttgttatttt	cccatctctc	ttaaaaaagt	960
ctttatgatt	attttatata	gttttgagaa	cttttaaaggc	actttttttt	aaccttacat	1020
ttgcataaaa	atgttttagct	tttaagtaga	gagcaaatat	tgatcatata	ttttgatatt	1080
catgacctgt	ttgactatag	gagttttttt	taaaaaaatg	cactttgggt	ataaaaccat	1140
ggatgatttg	atccataaga	tttaaatgtg	ccaccattat	agtattccta	gacatgagct	1200
tgatgaatgg	tattctgtaa	ttataacgtg	ccccacatta	ttgtgtctta	attgccctta	1260
gcctgaattt	taatgatcaa	tttgttattg	ttgcagatgt	gaatattgtg	cataaaccta	1320
ctaaatttat	gtaaaattgt	ataaaataga	attagaagtc	actaagttct	ttctgtgtag	1380
aagtaataaa	tttattgtaa	cacaaaaaaa	aaaaaaaaaa	aaa		1423

<210> 52

<211> 1364

<212> DNA

<213> Homo sapiens

<400> 52

tctacagtaa	accccacat	taccttctag	ttggcataaa	aacaagaacc	acaaaaactt	60
gaaaaatctg	aaacagagaa	cagaaggcca	agagccagct	tctgcgttct	caactttatt	120
caacattaga	cttaccttat	ttctttcagt	gtgtagggac	aagatgtact	gctgtgtgtg	180
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtm	tataccttcc	tatccatttg	caagttaacc	240
tccatctggg	ttatttgccc	atgctatgct	ttcttcccat	tccctgctg	tctattctaa	300
gccccagac	tcaagcctct	agactcttgg	atgaaacagt	gagaagaaaa	cattttctga	360
cttacccttt	tggaatctcc	tccattatta	cccaggcttt	gctttaagtt	gcactttaaa	420
tcacactgtc	ctattaatgc	gatctggcat	cttctcccac	aagcccccta	caggaacaaa	480
ctaccccccta	ccttaactct	aaatggttct	ttagactata	gtctgtctcc	tctgacctga	540
aatcctcttt	taggcaatag	gccgactttt	agaagcagcc	aggctctggg	agaaatgggc	600
ccccatacca	actgtggact	tggaaatata	gcagagtagt	aggcacagtt	gtaaaaaggg	660
gagatttcgg	taggcacagg	tgtaaaaagg	ggagatttca	gaagacgggc	gaaacaactg	720
atgggggaag	atacctgggg	tgagaagatg	agaaagaaat	gatgctgagg	gacctagtga	780
aatcaatgaa	actcttgagt	cttgcttagg	ctcgcaaaaca	agaagtgggg	aggcttggga	840

29

aattaggata	tgacatatat	gaaagggttta	tctaagaaga	aagaaacaga	gataatatat	900
atatagaaag	atgtatatag	atgtatatat	gccccaaatat	attgaatgta	cagaaaggaa	960
atattcagag	acttttagcat	tgggggcaga	tatcttgccc	tggttatggg	gtaggacacc	1020
cagattttca	gacttacaat	cagtggctct	ggtttcaaca	tggaaagtga	atagctgatg	1080
aaggatttca	agtctggaaa	tgggaaagg	aggggtagaa	gttccttttc	agaaataaag	1140
aggcattcat	gaaggctttt	gttgaatcct	atgctactga	taggaccagg	aagagaggaa	1200
ctcggaggac	aataggagg	ggaagtcttg	gaaaatctac	cacttacact	gtgtgtcccc	1260
atccccagca	gcgtcctgcc	actgtagcgc	ctttttaaaa	taaataaaat	aaaataaagc	1320
accaaaaaaa	aaaattaaaa	aaaactggag	ggggggcccc	gtac		1364

<210> 53

<211> 2288

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (940)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1279)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1798)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2280)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2285)

<223> n equals a,t,g, or c

<400> 53

gatcccatto	ttctctcgtt	gggaatgctt	gtggggggaa	aaagaaacgc	aatagataaa	60
gcggggcgca	tgcgctcccg	gcacaggcwt	cgattgtgag	gaargccggc	tagtctccga	120
gctcatcccg	ccttgcgcat	gcggagaagg	taaaccagcg	ccccgagttg	aggcgcgggt	180
ttggtggcgc	gtttcagcga	agtcgcacgt	gaaggatagc	agtggcctga	gaaagaccca	240
gtcatggcag	cctccagcat	cagttcacca	tggggaaagc	atgtgttcaa	agccattctg	300
atggtcctag	tggcccttat	cctcctccac	tcagcattgg	cccagtcctg	tcgagacttt	360
gcaccaccag	gccaacagaa	gagagaagcc	ccagttgatg	tcttgaccca	gataggctga	420
tctgtgcgag	ggacactgga	tgccctggatt	gggccagaga	ccatgcacct	ggtgtcagag	480
tcttcgtccc	aagtgttggt	ggccatctca	tcagccattt	ctgtggcctt	ctttgctctg	540
tctgggatcg	ccgcacagct	gctgaatgcc	ttgggactag	ctgggtgatta	cctcgcccag	600
ggcctgaagc	tcagccctgg	ccaggctccag	accttcctgc	tgtggggagc	agggggcctg	660
gtcgtctact	ggctgctgtc	tctgctcttc	ggcttggtct	tggccttgct	ggggcggtatc	720
ctgtggggcc	tgaagcttgt	catcttctctg	gccggcttcg	tggccctgat	gaggtcggtg	780
cctgaccctt	ccaccggggc	cctgctactc	ctggccttgc	tgatcctcta	cgccctgctg	840
agccggytca	ctggctcccg	agcctctggg	gcccactcgc	aggccaaggt	gcgagggctg	900
gaacgccagg	tggaggagct	gcgctggcgc	cagaggcagn	cggccaaagg	ggcccgagct	960
gtggaggagg	agtgaagcgg	atgccccaca	caccgcccag	gtcataccaa	agagctgagc	1020
tgcttcgggg	ccatgcagcc	ctcctgccag	ccccctgccc	ttttcttgcc	ctgtctctga	1080

30

accttcagaa	cattgatecct	tgccgcagcc	ccactagcca	agagaaacag	agaaagacca	1140
ttccccctgc	ctgtccttgc	ggccctgtct	tctgaggttc	tctgtctggg	gttggctctc	1200
ttaacccttt	ctctgctccc	agcctgcctc	accagggaag	gttggagggg	cctccctctg	1260
gcttctgcat	ctgcgccana	aacatcactg	ccgttggtct	ctcatgactt	aactggcttc	1320
cctctgctgc	tgcttggct	tcctccta	gtcgtgctc	tcctgtcctt	ctgaagttgc	1380
tccttggcca	aatctccagc	tcccttcttg	ttttcctcat	cctcctacce	tgtactccca	1440
ccaaaccatg	gtcctttaag	gcacgctcct	gtcctcctca	ttgccagca	gtagggaggg	1500
gcaggggtaa	ggggacctga	ggataaagg	tggggaaaca	gggtcccctg	aggcctgtgg	1560
gggctgcagg	ggaggaggat	gtacctgtg	tctctttcaa	gtgccttaat	ccgagccagc	1620
agggccttct	gcttgccctgc	tgccatactg	tatgtaggaa	agtgttctgt	ggctgctttg	1680
tgtcaagaaa	agagcagtca	ctctcagaat	cttgattccc	catcagccaa	agcaaaagat	1740
ggctgctgct	ttgtaggcat	gtgcctgcaa	gtgggacctt	gctgggcatt	atatgccntg	1800
tgggggtttc	agagaccctg	aaagaggagg	gaggaccctc	ctccttgtct	gcacaactgc	1860
atgcacttct	ctccccatcg	ctccacaacc	tgaaccgag	aaggagtgtc	tgaccagtgc	1920
ccaccccggc	agcccgagg	gaacacaggc	agctccttct	ccttcacgtg	gtctgcagag	1980
agcaggggtga	gctgccagct	gcccctctcc	accagggtac	cctgtcttgg	tggtagggg	2040
ccacttttcc	tttgaggctc	tagtggaggt	ggatgtcctt	ctctgccagg	cttggcacat	2100
gatgtgaaga	ataaatgccc	aattcttact	gttcaggttt	gatgtggaat	cacagctgca	2160
gtgatataata	ttttttatca	gtgcttggtt	ggtttttaaat	aaagtgcacg	ctattttatt	2220
atcttgttct	gaataaaatg	tatttactcc	aaaaaaaaaa	aaaaaaggs	ggccctctan	2280
agggncca						2288

<210> 54

<211> 1512

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (2)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (8)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (16)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (21)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (29)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (528)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (600)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1496)

<223> n equals a,t,g, or c

<400> 54

cngaaaaancc	ccgtgncatt	ntgggaaana	acgccccgca	ggtaccggtc	cggaattccc	60
gggtcgaccc	acgcgtccgc	ccacgcgtcc	gctcgctggg	ctttgtcctc	ttctgtgatg	120
aagttagaca	gtggtacgta	aatgggggtga	attatttttac	tgacctgtgg	aatgtgatgg	180
acacgctggg	gcttttttac	ttcatagcag	gaattgtatt	tcggctccac	tcttctaata	240
aaagctcttt	gtattctgga	cgagtcattt	tctgtctgga	ctacattatt	ttcactctaa	300
gattgatcca	catttttact	gtaagcagaa	acttaggacc	caagattata	atgctgcaga	360
ggatgctgat	cgatgtgtgc	tycttcctgt	tcctctttgc	gggtgtggatg	gtggcctttg	420
gcgtkgccar	gcaagggatc	cttaggcaga	atgagcagcg	ctggagggtgg	atattccggt	480
cggtcatcta	cgagccctam	ctggccatgt	tcggccaggt	gcccagtnac	gtggatggta	540
ccacgtatga	ctttgcccac	tgacacctca	ctgggaatga	gtccaagcca	ctgtgtgtgn	600
agctggatga	gcacaacctg	ccccggttcc	ccgagtggat	caccatcccc	ctgggtgtgca	660
tctacatggt	atccaccaac	atcctgctgg	tcaacctgct	ggtcgccatg	tttggtctaca	720
cggtgggcac	cgctccaggag	aacaatgacc	aggctctgaa	gttccagagg	tacttcctgg	780
tgcaggagta	ctgcagccgc	ctcaatatcc	ccttcccctt	catcgtcttc	gcttacttct	840
acatggtggt	gaagaagtgc	ttcaagtgtt	gctgcaagga	graaaacmtg	gagtcttctg	900
tctgctgttc	aaaaatgrag	acaatgagac	tctggcatgg	gagggtgtca	tgaaggaaac	960
taccttgtca	agatcaaac	aaagccaacg	acacctcaga	ggaaatgagg	catcgattta	1020
gacaactgga	tacaaagctt	aatgatctca	agggctctct	gaaagagatt	gctaataaaa	1080
tcaataaaaa	ctgtatgaac	tctaattggag	aaaaatctaa	ttatagcaag	atcatattaa	1140
ggaatgctga	tgaacaattt	tsctatcgac	tactaaatga	gagattttca	gaccctctgg	1200
tacatgggtg	atgattttta	atcaccctag	tgtgctgaga	ccttgagaat	aaagtgtgtg	1260
attggtttca	tacttgaaga	cggatataaa	ggaagaatat	ttcctttatg	tgtttctcca	1320
gaatggtgcc	tgtttctctc	tgtgtctcaa	tgccctggac	tggagggttg	tagtttaagt	1380
gtgttcttac	cgctcctttt	ttcctttaat	cttatttttg	atgaacacat	atataaggaga	1440
acatctatcc	tatgaataag	aacctggtca	tgctttaaaa	aaaaaaaaaa	aaaaanaaaa	1500
aaggcgccgc	gc					1512

<210> 55

<211> 1357

<212> DNA

<213> Homo sapiens

<400> 55

ggcacgagtt	tatttacagg	catataaaat	gaaattgtga	gatgttttgc	aagcttcttt	60
ttactttgag	tagcttttaa	tttgtatgtt	tttatgtgga	tgaagagcat	tttttatgct	120
tttgtgcaat	aggttccaat	atgcatttat	tagacatctg	tttaaattgg	aatgtagcat	180
ttattttgct	aaattgaaag	ggaacataga	tggaaattcca	aaatatgtac	attcagctgt	240
ttgggttttc	gtttttcatt	gttattattg	tgagaatgct	gttattgggg	ttgtgtgtga	300
gtgcccgtca	gccagtgatg	cctcgggcca	cgctgtgggg	ccacctcagt	cctgcctggg	360
tcctgggtgcc	ttggacccca	cgctgctgtg	gccaggctgc	ccctggggcg	ggccatgtgg	420
cctcagacca	caagagcgga	ctgccttggc	ccaagcactg	cagctgcctg	cacccccggg	480
cttcgcagcc	ttgcttggtt	tctctgaaca	gcaacagaac	agtgttcaca	gcgattcaaa	540
gggtggcatt	gggttgagcg	ttctgggtac	aagccaacct	agtcccacgt	tgtacgtgaa	600
tgtttaaatg	gctctcaaaa	catggaaaat	aagttagtg	cacatagcta	aatcacaaaa	660
catccaattt	ctctgtttcc	tcaggaaagtc	attactgcgc	caccacatca	catgacctta	720
acatgatcaa	tgtattttctc	tgccctgaca	tttaaataca	taaattgaga	taagttagatt	780
agaaaatcat	tcaaatgata	ccataatttg	tacgggacag	gggtgcgggca	atggccacgt	840
ggccaaggcc	ccgcaggaac	gcgccgaggt	ctccctcacc	ctccagggtg	ccttcgcacc	900
caacagtgcg	tctgaggaac	gagctgcagt	ttgagcgttc	ccctgagatg	tgcgtagcct	960
ccgtgtaaat	gtccactccc	atggcttaat	tgccctatcag	acgcattttc	ccagacgaaa	1020

32

gcaatgttgg	gttggggaag	acagtgcagc	cacccagcct	ttaccagcag	cgtacggcag	1080
acgaaggcag	tcgaggtgtg	gaggtgatca	cgaagataca	tgtgtttgac	tgtttaattt	1140
gaaagtttac	atTTTTtatg	ctttgtgttg	gtgtgtaatt	tttgtactct	tggtggctag	1200
tttttgtcaa	atctTTTTtg	gaatattgct	taaatgtttt	gattttatga	tagtgaagct	1260
tgtattcagt	gttttgccaa	ttaatatatt	atgcttgtaa	taaaagcaaa	agaaaagctt	1320
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaa			1357

<210> 56

<211> 1989

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (31)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (161)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (162)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1702)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1943)

<223> n equals a,t,g, or c

<400> 56

ttaaattgaaa	tcaaaattgg	ccatttgaca	naagttgggt	tttccccctt	ctgcattttt	60
aggacctcaa	agtaattgtt	atccagaaac	tgctatcatt	accagggatt	cattcgtgta	120
tttaacaaca	tggggcatac	atTTTtgcca	aatttgaaaa	nntcttaaca	tacaccccaa	180
aatccctgcc	ccaaatttaa	gaactagggt	ggacacagtg	cgtttttcca	tgctgcactc	240
tctgtgatgg	ggctacgata	cgtgggagca	gagaatgggg	aggttggagc	gcatgccaga	300
tgaggatcta	tcagcaatgg	gacgggkcct	ccactttagc	atctcyacce	tgctcctytc	360
agaggaccgc	ctttcattgc	attcagctgt	gatggttagc	cgaacacagg	tgccaccgag	420
acgaggagag	caggagcctt	gtgctctctc	tgcatctgag	gcaggacagc	acagggtayg	480
gagcagtctg	cagagaggcc	agctcatcag	ggaagcactt	gtcttccacc	ttgggctttg	540
actgagcact	gggcaattgg	mcycctggga	tcaaygaaat	aatcctaarc	agagttactc	600
tatgtcacac	tatggaatgt	tccaagtasr	tggccgtggt	ttcaaaagat	rtattttctc	660
cttttgttgt	tgccatttca	taggttttag	attgggtgtg	tgktctctct	ctctgaatgg	720
cactcraatg	tttgctgact	cctactctgt	gtgactgggg	tgtacagcta	tggaactgat	780
catcccatcc	catcatcttt	catgatcaaa	gcagtctctt	cttttttgac	agctgaagaa	840
gcatcggtag	ggaatccaga	aggagcgttc	atgaagggtg	tacaagcccg	gaagaactam	900
acaagcactg	agctgattgt	tgagccagag	gagccctcag	acagcagtg	catcaacttg	960
tcaggctttg	ggagttagca	gctagacacc	aatgacgaga	gtgatkttat	cagtacacta	1020
agttacatct	tgccwtatTT	ctcagcrgta	aacctagatg	tgraatcamt	gttactaccg	1080
ttcattaaac	tgccaaccmc	aggaaacagc	ctggcaaaga	ttcaaactgt	aggccaaaac	1140
crgcararag	tgaakagagt	cctcatgggc	ccaaggagca	tccagaaaag	gcacttcaaa	1200
gaggtrggaa	ggcagagcat	caggagggaa	cagggtgccc	aggcatctgt	ggagaacgct	1260

33

gccgaagaaa	aaaggctcgg	gagtcacgcc	ccaagggags	tggaaacagcc	ycacacacag	1320
caggggcctg	agaagttagc	gggaaacgcc	rtctacacca	agccttcstt	caccaagag	1380
cataaggcag	cagtctctgt	gctgamaccc	ttctccaagg	gcgcgccttc	tacctccagc	1440
cctgcaaaa	ccctaccaca	ggtgagagac	agatggaaag	acwwmacmca	crctatttcc	1500
atthttagaaa	gtgcaaaggc	tagagttaca	aatatgaagg	cttctaaacc	aatttcacat	1560
tccagaaaaa	aataccgctt	tcacaaaact	cgctcccgcg	tgacccacag	aacacccaag	1620
gtcaaaaaa	gtccaaagtt	cagaaagaaa	agttatctga	gtagactgat	gctcgcaaac	1680
aggcctccgt	tctctgcagc	gnagagcctc	ataaattccc	cttcacaagg	ggctttttca	1740
tccttaggag	acctgagtc	tcaagaaaac	ccttttytgg	ragtatctgc	tccttcagaa	1800
cattttatag	aaaccactaa	tataaaagac	acaactgcaa	gaaatgcctt	ggaagaaaat	1860
gtttttatgg	aaaacactaa	catgccagaa	gtcaccatct	ctgaaaacac	aaactacaat	1920
catcctcctg	aggcagattc	cgntgggact	gcattcaact	tagggccaac	tgtaaacaca	1980
actgagaca						1989

<210> 57

<211> 2543

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (2538)

<223> n equals a,t,g, or c

<400> 57

ctccgttggg	aacttgggct	gagtaccgcg	gcggggcgcg	gcragggcgcc	ctagacatct	60
tctccctccc	ttgcctcaga	tttattgcta	aacatgggtg	catttttggg	taaacccaaa	120
actgaaaaac	ataatgctca	tggtgctggg	aatgggtttac	gttatggcct	gagcagcatg	180
caaggatgga	gagtggaaat	ggaagatgca	cacacagctg	ttgtaggtat	tcctcacggc	240
ttggaagact	ggtcattttt	tgcagtttat	gatggtcacg	ctggatcccc	agtggcaaat	300
tactgtctca	cacattttat	agaacacatc	actactaacg	aagacttttag	ggcagctgga	360
aaatcaggat	ctgctcttga	gcttttcagt	gaaaatgtta	agaatgggtat	cagaactgga	420
tttttgaaaa	ttgatgaata	catgcgtaac	ttttcagacc	tcagaaacgg	gatggacagg	480
agtggttcaa	ctgcagtggg	agttatgatt	tcacctaacg	atatctactt	tatcaactgt	540
ggtgattcac	gtgctgttct	gtataggaa	ggacaagtct	gcttttctac	ccaggatcac	600
aaaccttgca	atccaaggga	aaaggagcga	atccaaaatg	caggaggcag	cgtgatgata	660
caacgtgtta	atgggttcatt	agcagtatct	cgtgctctgg	gggactatga	ttacaagtgt	720
gttgatggca	agggcccaac	agaacaactt	gtttctccag	agcctgaggt	ttatgraatt	780
ttaagagcag	aagaggatga	atthttatct	ttggcttctg	atgggatctg	ggatgttatg	840
agtaatgagg	agctctgtga	atatgttaaa	tctaggcttg	aggtatctga	tgacctggaa	900
aatgtgtgca	attgggtagt	ggacacttgt	ttacacaagg	gaagtcgaga	taacatgagt	960
attgtactag	tttgcttttc	aaatgctccc	aaggtctcag	atgaagcggg	gaaaaaagat	1020
tcagagttgg	ataagcactt	ggaatcacgg	gttgaagaga	ttatggagaa	gtctggcgag	1080
gaaggaatgc	ctgatcttgc	ccatgtcatg	cgcatcttgt	ctgcagaaaa	tatcccaaat	1140
ttgcctcctg	ggggaggtct	tgctggcaas	cgtaatgtta	ttgaagctgt	ttatagtaga	1200
ctgaatccac	atagagaaag	tgatgggggt	gctggagatc	tagaagaccc	atggtagcct	1260
taaaaacctt	ctaaaatgct	tttrattctg	aaaattgggg	gaaaaaactt	ttaatcacia	1320
ttttcttcaa	tacaagggga	aaatattctt	gcggattccc	aacgttttgt	gatatgagca	1380
gaaaatcatt	agcatttccc	atcatttgtt	catatttgtg	ttttctgaca	gttgccactt	1440
gtagcattgc	ctgtactaca	gtattttttg	ccaacctcag	gcatactcgt	tacatctgta	1500
ttgaactttc	ggccctagaa	accagtggag	ttattttacc	acaaatcaac	aatgtgcctg	1560
aggtgcatgg	gaaatatagt	tagctatact	atgaaaatac	attatgtttt	ttttctttta	1620
acaaaacaca	caacatgtaa	gcattgtaaga	gtaaagaatt	gtatgatgat	ttcctttttt	1680
cagttcacca	agttggaagc	cttttgcagc	tctgtggcct	ggaatttcac	ttgagcaatt	1740
tctataggat	atgtattttat	tattgattgt	tatttaawww	wwttccamtt	ttacctgtat	1800
taccaaactg	ggttctccaa	taatgtccaa	attgtaattgt	tgcccttgctt	caagataaag	1860
tgtatttggg	aataatatta	taaacccttm	caaattttat	gcattgtatct	actgcatcct	1920
tcaactctca	ctagaaaatc	ttttgaaacc	aaatggatta	atttatggct	atttataatt	1980
tgctttgaca	tctcactgtt	ggaaattttt	taaagatgag	atttgccttt	ataatgtaaa	2040

34

ttgtgatttt	tgttttacat	gtgggtttct	atagttttta	ttttttcagc	ttttaagata	2100
cgaagttttgt	gtaatttggt	atttttaatc	attttatgta	ttttaaaagc	tcagaatata	2160
acattgaaat	tactataaat	acatttaaaa	ttatctat	tagatctaag	gaaatactac	2220
agagatattt	tcatgggttc	agtaactttt	cattttataa	cattgggcac	ggtacagagt	2280
gattgtcaca	taaggtactt	gaagatttat	tagtttaatt	ctatttttac	agtaaccttg	2340
aattcttctg	agttttgcat	gtattaaatt	caattaatgc	tgaacatgaa	gagtaaagta	2400
tttatctgaa	agaagtttct	gggttaggag	aagtaatgaa	tgtatccatt	tgtacatggt	2460
ttacatgttg	tgatgctttt	gtaaacattt	tcctgtatgt	ttaaattgtg	tttcagcagg	2520
atgtagttgc	ccttgtgnag	gtt				2543

<210> 58

<211> 777

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (766)

<223> n equals a,t,g, or c

<400> 58

ggcagagcgt	taagtcctca	ttccctcctt	ctcctcttcc	tttctctgca	gtaggggagg	60
cccactcccg	ksggatctat	cttgggatcc	catggcttcc	tttactgggc	tctggggccc	120
cttcacctgt	gtaagcagag	tgctgagcca	tcactgttcc	agcaccactg	ggagtctgag	180
tgcgattcag	aagatgacgc	gggtacgagt	ggtggacaac	agtgccctgg	ggaacagccc	240
ataccatcgg	gctcctcgct	gcatccatgt	ctataagaag	aatggagtgg	gcaagggtggg	300
cgaccagata	ctactggcca	tcaagggaca	gaagaaaaag	gcgctcattg	tggggcaactg	360
catgcctggc	ccccgaatga	ccccagatt	ygactccaac	aacgtggtcc	tcattgagga	420
caacgggaac	cctgtgggga	cacgaattaa	gacacccatc	cccaccagcc	tgcgcaagcg	480
ggaaggcgag	tattccaagg	tgctggccat	tgctcagaac	tttgtgtgag	ttgagcccag	540
gcctctgggt	gcaggactcg	tgaatggagc	agttctgaga	accacccttt	tgctaaggga	600
gcttgggagc	cacatggctg	ctcccttcac	actgggtaac	agtgtagtat	cctgtgagag	660
aaataatgta	ttcatttatg	tgtttttcca	gagctttctg	ggatgtggga	aaataaatta	720
cactgaagca	gttgaaagg	gaaaaaaaaa	aaaaaaaaaa	aaaaanaaaa	actcgag	777

<210> 59

<211> 879

<212> DNA

<213> Homo sapiens

<400> 59

gctgcatgct	ggcggggaac	taggaagcct	ccccaacctc	tggccccgtg	gagccctcag	60
cctcagctgc	agtggaggca	cctcgggctc	tggggcaacc	aagtgtgaca	ggtggctgtg	120
cacgggcaga	ggtcctgtgg	aagatttcat	gtgacgggca	gaagaggagg	aggaggcagg	180
ggaggaagca	catccatgaa	cagggtgtgc	tgggggcagc	ctgggtggtc	gtgaaatagg	240
actcagtggc	cttgagtcct	catttaggcc	ctgatgttct	ttagcctgcc	tggccttttg	300
caaatcgcca	gcttcacgca	caacctcatt	tttcaccttt	gggtgtgggg	gtcagagtcg	360
ggagagcacc	tgcaaaagcca	caatgatcca	gacacacggc	aagggtggga	cattcccata	420
aggctcctcg	gggagagcag	cgcttctgtg	ccggggagca	gcgaaggcca	cacaggagga	480
ccgcacctc	ctcgtgtcgg	tggtccgct	ggtataatca	ggactcacgt	ggtgttcctc	540
gtgtcgtggc	ccttattgca	gaggagcag	cacaggcttt	cctggaagct	cccctcggtc	600
atgtgggggtg	actccagaga	ccccacctt	gcgagactgg	accagtccaa	gtggcctkga	660
gccacarcgg	cctkgcagta	cctkgggagg	gggtgatgac	aggtgcacac	ggaggcccat	720
gtggtctgtc	tggagaatgc	cggagatgtg	aaatatgtaa	tcctgagtgt	ggcttctaga	780
aggaagggtc	gcaaagctga	atatccactc	gtgctgttcc	cttctcacag	gagattcctg	840
tcaacgtccg	attctgcctc	gaaggcagga	ggagtaagg			879

<210> 60

<211> 1161

<212> DNA

<213> Homo sapiens

<400> 60

ggcagcagtc	gtccagcccg	cggcgagagc	gggtatgtgg	gcgggaggcc	ggagcagctg	60
tcaggctgaa	gtcctgcgag	cgacgcgcgg	cggggcggcg	agaggaaacg	cggcgcggg	120
ccgggccctg	gagatgggtc	ccgggcgcgc	gggctgggtg	tgtctcgtgc	tctggctccc	180
cgcgtgcgtc	gcggcccacg	gcttccgtat	ccatgattat	ttgtactttc	aagtgtctgag	240
tcctggggac	attcgataca	tcttcacagc	cacacctgcc	aaggactttg	gtggatatctt	300
tcacacaagg	tatgagcaga	ttcaccttgt	ccccgctgaa	cctccagagg	cctgcgggga	360
actcagcaac	ggttttcttca	tccaggacca	gattgctctg	gtggagaggg	ggggctgctc	420
cttcctctcc	aagactcggg	tgggtccagga	gcacggcggg	cgggcgggtga	tcattctctga	480
caacgcattg	acaatgacag	cttctacgtg	gagatgatcc	aggacagtac	ccagcgcaca	540
gctgacatcc	ccgcctctct	cctgctcggc	cgagacggct	acatgatccg	ccgctctctg	600
gaacagcatg	ggctgccatg	ggccatcatt	tccatcccag	tcaatgtcac	cagcatcccc	660
acctttgagc	tgctgcaacc	gccctggacc	ttctggtaga	agagtttgtc	ccacattcca	720
gccataagtg	actctgagct	gggaagggga	aaccaggaa	ttttgctact	tggaatttgg	780
agatagcatc	tggggacaag	tggagccagg	tagaggaaaa	gggtttgggc	gttgctaggc	840
tgaagggaag	gccacaccac	tggccttccc	ttccccaggg	cccccaaggg	tgtctcatgc	900
tacaagaaga	ggcaagagac	aggccccagg	gcttctggct	agaaccgaa	acaaaaggag	960
ctgaaggcag	gtggcctgag	agccatctgt	gacctgtcac	actcacctgg	ctccagcctc	1020
ccctaccag	ggctctctgca	cagtgcctt	cacagcagtt	gttggagtgg	tttaaagagc	1080
tggtgtttgg	ggactcaata	aacctcact	gactttttag	caataaagct	tctcatcagg	1140
gttaaaaaaa	aaaaaaaaaa	a				1161

<210> 61

<211> 687

<212> DNA

<213> Homo sapiens

<400> 61

ccgggtcgac	ccacgcgtcc	gactagttct	agatcgcgac	ggccgcctt	ttttttttt	60
tttactgcca	ggtagcaggc	tttattggga	agggacaaag	cctcaggagc	tgggtgcccc	120
agaggctgct	gggtcttgag	ccacagctgc	agccaatgca	gcagtcgcgc	ctccttcttc	180
cgtttctggt	tttcctcctt	gagggttgcg	ctccttcttc	tctaggtcct	ggagcagctc	240
ctggaagcgg	gcactccttg	ggccacctg	gtagcccagg	agctcctggg	cctcagcctg	300
cagtcggggc	ctcctctcct	tgtcagcctg	ggccttctcc	cagttctccc	gctgctgctg	360
ctgccagttc	acaatcatct	gtggcatctt	ggccatgcac	tctgcgatgt	gctgctccct	420
ctcccagcgc	ttctgtctct	cggccagctg	cttcacccgc	agcgactcct	gcatggctgc	480
caggctcggg	taccattcgc	gttcttcggc	ctccagctcc	cgcagctgct	ccggcgacgg	540
ccataacgaa	ccggggacca	ccccggaggc	ggcgccgtaa	cgcgcgaact	gcttagcgcg	600
tagcgcggtc	ccagctgcca	ccgcggggtc	aggaggctct	cggggtcttg	ccaccggggg	660
cccgggtcgg	cgcgggggcg	gccgctc				687

<210> 62

<211> 518

<212> DNA

<213> Homo sapiens

<400> 62

acgcgtccga	gatacattcc	atgaatacct	agttttattga	gagtttttag	catgaaggac	60
tgctgaatth	tgtaaaagc	tttttctgca	tctattgaga	taatcatgtg	gtttttgtct	120
ttggttctgt	ttatgtgatg	gactatgttt	attgatttgc	atatgttgaa	ccagccttgc	180
atctcaggga	tgaagccaac	tcgatcggtg	tggataagct	ttttgatgtg	ctgctggatt	240
tggtttgcca	atattttatt	gaggattttt	gcatacagtg	tcttcaggga	tattggctta	300
aaattctctt	ttttttgttg	tgtctctgcc	aggcttttgt	atcaggatga	tgctggcctc	360
ataaatgagt	tagggaggat	tccctctttc	tattgatcag	aatagtttca	gaaggaatgg	420
taccagctct	tctttgtacc	tctggtagaa	tttggtgtg	aatctatctt	gtcctggaat	480
atthttgggg	ttggaactca	aaaaaaaaa	aaaaaaaaa			518

<210> 63
 <211> 911
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (911)
 <223> n equals a,t,g, or c

<400> 63
 gtctgaccag gggactaaa taaaccggcc ctaacacttc catctccacc caccatct 60
 ccctggcgat gtgtccagc ccaagcagcc tccgtaggct ttagatcctg tgggtgccag 120
 atccagtcct ttctaatacc ctgagtcaac acattactcc tgcaggctctt aggctacaat 180
 gcaggctccct tgaggggccac caacatggag gtaggcagtt tctaggactg tccccagtac 240
 atctcaccac ccacagccct ttttttgccct tgattcgagc ctcaccctgg ccttttggtc 300
 tccccctgct gagagagacc tgaggagggg acagagccca gcccctctcc tgtggctgag 360
 caggcctctg tgtccatgac acctgtcttc cgggctggg ggtgtgggt gtatgtcctc 420
 cctactggct tccccggccc ctgctgcatg atgctcttg aactcttccc caaggagtca 480
 gtccccagg cctatcaggg gatccttttg tatctgact ttgggtttta gtttcaaagc 540
 tccatcagg acagcttgca tttcaggatg tgtggaaagc tcgggtgagg gctgccctgg 600
 ttcacatag ctccaccttc ctcggaagga gtgggtgtt ggagacccc catccatggc 660
 aactagctc agcactgcat ttcccagat gattcccaag acagctggtg cctcctggct 720
 ttctgtgcc agccaaggg gcaccacaga ggacctgga tcctttgcct cttcttggtt 780
 gaaggatctc tatgtatgtg tgtatataaa tatagttttt tatctatata tataaaaaaa 840
 aaaaaaaaaa aaaaaaaact cgaggggggg cccggtaccc aattcgccat atggtgatgg 900
 caaatgggaa n 911

<210> 64
 <211> 963
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (2)
 <223> n equals a,t,g, or c

<400> 64
 tncagaggcc ctgcggagtt gttcagaacc ccaactctct ctggctggct acccctgaa 60
 ctactgggtc tctggaccca ttgtgccag ccaccccaa aagccctcag gcgagagctg 120
 cctgaggagg caccgtgag gaggaaagga gaaagattga agttccaagt gagattgaga 180
 gatctcccta gaggcagctg aagaggagaa gtcccgcatc agcctcatcc caccagaaga 240
 acggtggtaa gcggccaggc tccgtggrag ccagggccca magcccttgg ccagktkgtg 300
 gaaacagctg ctgggatggg tatgccctt gtcactgtca cagctgccac cttccctact 360
 ctctcatgtc ctctagggc ctggcctgag gtggaggcgc cagaagctcc tgcattgccc 420
 gtggtgcctg aactccctga ggtgccatg gagatgcctt tgggtgctgcc ccagagctc 480
 gagctgctct cactggaagc agtgcacagg taccaggrag gtggcacctt gatgggggtg 540
 accggggctg aagcctctgc taatggttct tgatccctat agggcagtg cactggagy 600
 gcaggctaac agggagcccg acttcagcag cctggtgtca mctctcagcc cccgcaggat 660
 ggctgcccgg gtcttctamc tgcctctggg tgartgtatg catgtgtgtg tgtgtatgtk 720
 gggcagggac acagagacca gaggccgta cagggactcc cccgacctgc cctctcctcg 780
 cctcttgacc agtgccttca gcgaacaga ttcttcacgt gaaacaagaa aagccatatg 840
 gtcgcctcct gatccagccg gggccagat tccactgagg ttagagtcca ttacaaaagc 900
 tgccaggaaa ccggccactt ctagtaaac acgtcgtgcc tcactgaaaa aaaaaaaaaa 960
 agg 963

<210> 65

<211> 1001

<212> DNA

<213> Homo sapiens

<400> 65

cctactctc	atctgctcca	gccccctgac	cttatagttg	cccagctttc	ctggcaattg	60
actttgcccc	tcaatacaca	ggatttagca	tccagggaag	atgtcggagc	ctcagatgtt	120
aattttctaa	ttgagaatgt	tggcgctgtc	cgaacctgga	gacagagtat	cagcgccttt	180
gcttgctgct	gtttttgctg	ttttttgatg	ctgggaacca	ccacctaaag	atagtaaaga	240
aaacacagga	agctttcccg	aaaacaaaaa	gtcctttctc	ctgattcacc	aaaaaataaa	300
atactgacta	ccatcactgt	gatgagattc	ctatagtctc	aggractgaa	gtctttaaac	360
aaccagggac	cctctgcccc	tagaataagr	acatactaga	agtcccttct	gctaggacaa	420
cgaggatcat	gggagaccac	ctggaccttc	tcctaggagt	ggtgctcatg	gccggctctg	480
tgtttggaat	tccttcctgc	tcctttgatg	gccgaatagc	cttttatcgt	ttctgcaacc	540
tcacccaggt	ccccaggtc	ctcaacacca	ctgagaggct	cctgctgagc	ttcaactata	600
tcaggacagt	cactgcttca	tccttccctt	ttctggaaca	gctgcagctg	ctggagctcg	660
ggagccagta	taccccttg	actattgaca	aggaggcctt	cagaaacctg	cccaacctta	720
gaatccttga	cctgggaagt	agtaagatat	acttcttgca	tccagatgct	tttcagggac	780
tgttocattc	gtttgaactt	agactgtatt	tctgtggtct	ctctgatgct	gtattgaaag	840
atgggtat	cagaaattta	ctcgcttgga	tctatccaaa	aatcagattc		900
gtagccttta	ccttcacctc	tcatttgggg	agttgaattc	cttaaagtcc	atagattttt	960
cctccaacca	aatattcctt	gtatgtgaac	atgagctcga	g		1001

<210> 66

<211> 1558

<212> DNA

<213> Homo sapiens

<400> 66

gcacatgagg	ccttgacagt	ctccttacgc	acatgcgggc	cttgtagctc	tccttaccca	60
catgcggggc	ttgccgtctc	ccttaccac	atgtgggcct	tgcagctctc	cttaccacaca	120
tgcggccttg	cagctctcct	taccacatg	cggccttgca	gctctcctta	cccacatgag	180
ggccttgccg	ctctccttac	ccacatgggg	ccttgccgct	ctccttacc	acatgggggt	240
cttgacagt	tccttacgca	catgcggggc	ttgcagctct	ccttaccac	atggggcctt	300
gcagctctcc	ttaccacacat	gcggccttgc	agctctcctt	acccacatgc	gggccttgca	360
tgctgttggc	tctggagcct	ctcgtctcac	aggtctctac	aggtgcaggc	cactcaccgt	420
ctggtggtca	ggaccataaa	ggacagggtt	atgttaaagg	ttttgcctca	aaccagaagg	480
cgaggaccct	ttctgtccag	ttgccggaat	gatgtcatga	ggaactgtgt	gcccaggcac	540
gctgtgctag	ttacaacatg	tgtttttgtt	tcattcccca	cacactgtaa	ggtgggcac	600
actggggcca	tcacacaggt	gaaacagaag	cccggaatc	actcgtcccc	ttgccagctc	660
atacaactag	tagccaaggc	agaatttgaa	ctcatgttgc	cctcagctcc	aaaacctgtg	720
tacttaaccc	ttgttctctc	ctgctggtgt	ctgtgtgatg	tcccatgtct	gtctgtctct	780
ctctaaagg	acagtgcac	accaggagga	taccagatg	ctggggggcc	ttgggacaga	840
gtctgggagg	attgagtga	ggagcagggt	aggggtgagc	tggagagaga	acgccctggt	900
ggagagtta	tgtagaaagg	ggattagggt	tccgggagga	accggatcca	tgtggtctgc	960
tgagatggct	gagtctggca	ttcagatgtg	ccaccaaca	gaagaggccc	tggaggggacg	1020
ccccctttgc	tgggtggcag	ccgtgggatt	ccggggtctg	ccttgagggt	cctggagagg	1080
atgtcgtggc	cctggcccta	gactcaagct	gcctgggtcc	agttcagccc	ggccactcct	1140
gctgtggg	ctagccagg	gccttcactc	caccgactgc	tgtgtgtttg	gtacatgggtg	1200
tcacccaggc	catgtgctta	gcaatgtgcc	tgacagccag	tgccgggtgtc	agccattaca	1260
gggacacacg	tgccctggag	ttgaggccac	gttctgtcac	ctaggcccgc	tcgtggctct	1320
gggctggg	aaacccccct	ttgaaaggat	tcctttttgc	ccctggcata	ggctctcatt	1380
gtcctagtga	acagctacat	ctttttaaca	agccagaaaa	ggccagctgg	cagtggtctt	1440
gcctgaaatc	ccaagactgg	ctggccgaag	caggaggatc	acttgaggcc	agcctggcca	1500
aagtaagcaa	gactctgtct	ctacaaaaaa	ataacaaaaa	aaaaaaaaa	aactcgag	1558

<210> 67

<211> 1322

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (11)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (690)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (719)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (720)

<223> n equals a,t,g, or c

<400> 67

ctragcaact	nagtgggac	ccccgrrct	ggcaggaatt	cggcacgagg	tggaatctgt	60
gaccagaag	taacaaactc	ctttcttga	gagcagttag	gtattccatt	ccttagtcca	120
tatgccaaca	tttaaaaagg	ccaaaaccag	aggcctagaa	aatgtatctg	gaagttgcag	180
tgagaccgtt	tttgatcatt	gtggccttcc	tggggctcag	tttctcgcct	ttgcaaatgc	240
cattttggca	gggatctgct	gtggggcatc	tccgtgcagg	tgagagctga	gttgcgcatc	300
tttctcaggc	tgcatcata	caggccccag	tgactcttgg	caggaggggg	cagccccctc	360
ctggatagcc	ccgsccaagg	ccgggargac	tgtgaagggg	ggatcccaact	gcctgacctc	420
agcctgtcgg	gccccacagc	gcgtctctct	gtggactggt	cggcggtctc	ctgtggcctg	480
tgtgtcctcc	gagtggctgg	agctttggaa	ccctattctg	tagcttggag	ctcctgagcc	540
tcaaagggca	ggggcctggt	tccttgccca	tccttgccca	gcctgatggc	ctgtgcttgt	600
ggactgtaca	tgggcacctg	ctttaacacc	tgaggagta	ggggctacca	agaagcatgt	660
ggctctgggc	ctccctggga	gagtcactcn	gcggcaggag	aactgagtg	gacacatcnn	720
ggagtgtctg	tctcatggac	scctkttggc	ctgcagcctg	gagagggggc	ctgaagtgtg	780
tgttccatgc	tcttgacccc	cagkaagcac	tcgcctctgt	tgaaagcctc	gtgccgcaga	840
gcgcgattgc	tgtcccgggt	ggacggccat	cacgggctcc	ttgctccggc	gatgccagcg	900
ctcctggtgt	tgtgtgtggc	tggtccccc	ttgtctcagc	cctgggctta	gagcaggcca	960
ggtgtcagag	cagtggtttt	gttgcttgaa	ggggggtgtg	tacctggctg	cagcctgtgg	1020
agagcgtgag	tggtgtgcag	aagcaaaggc	gggctcttgg	gagatgagag	caggcagccg	1080
gcctggaggc	ttccatgggg	ctggtcagct	cctgctggct	cttcctggca	caaatccacg	1140
ctgggctggg	tgccctggct	cacacctgta	atcccagcac	tttgggagcc	tgaagtggga	1200
ggatcgcttg	agcctgggag	ttcaagacca	gcctgggcaa	cactgtgaga	ccctatctct	1260
atTTTTAAAA	ataaaaaatat	aaagataacc	cttcctycaa	aaaaaaaaaa	aaaaaactcg	1320
ag						1322

<210> 68

<211> 865

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (445)

<223> n equals a,t,g, or c

<400> 68

gaattcggca	cgagcagacc	tgggctcgag	accataactg	tttggcttta	acagtacgtg	60
------------	------------	------------	------------	------------	------------	----

39

ggcggccgga	atccgggagt	ccggtgaccc	gggctgtggt	ctagcataaa	ggcggascca	120
gaagaagggg	cggggtatgg	gagaagcctc	cccacctgcc	cccgaaggc	ggcatctgct	180
ggtcctgctg	ctgctcctct	ctacctgtgt	gatccccctc	gctgcagctc	ctatccatga	240
tgctgacgcc	caagagagct	ccttgggtct	cacaggcctc	cagagcctac	tccaaggctt	300
cagccgactt	ttctgaaaag	gtaacctgct	tcggggcata	gacagcttat	tctctgcccc	360
catggacttc	cggggcctcc	ctgggaacta	ccacaaagag	gagaaccagg	agcaccarct	420
ggggaacaac	accytytcma	gccanctcca	gatcgacaag	gtaccagga	tggaggagaa	480
ggaggccctg	gtacctatcc	agaaggccac	ggacagcttc	cacacagaac	tccatccccg	540
ggtggccttc	tggatcatta	agctgccacg	gcggagggtc	caccaggatg	ccctggaggg	600
cggccactgg	ctcagcgaga	agcgacaccg	cctgcaggcc	atccgggatg	gactccgcaa	660
ggggacccac	aaggacgtcc	tagaagaggg	gaccgagagc	tcctcccact	ccaggctgtc	720
cccccgaaag	acccacttac	tgtacatcct	caggccctct	cggcagctgt	aggggtgggg	780
accggggagc	acctgcctgt	agccccatc	agaccctgcc	ccaagcacca	tatggaaata	840
aagttctttc	ttacatctaa	aaaaa				865

<210> 69

<211> 1150

<212> DNA

<213> Homo sapiens

<400> 69

gcggatcagg	agaaaataag	gaatgtcaaa	ggaaaagtaa	ttctgtcaat	gctggttgct	60
tcaactgtga	tcattgtgtt	ttgggaatth	atcaacagca	cagaaggctc	tttcttgagg	120
atatactact	caaaaaaccc	agaagttgat	gacagcagtg	ctcagaaggg	ctggttggtt	180
ctgagctggt	ttaacaatgg	gatccacaat	tatcaacaag	gggaagaaga	catagacaaa	240
gaaaaaggaa	gagaggagac	caaaggaagg	aaaatgacac	aacagagctt	cggctatggg	300
actggtttta	tccaaacttg	aaggaaatccg	aataactaaa	ctggactctg	gttttctgac	360
tcagtccttc	tagaagacct	ggactgagag	atcatgcggt	taaggagtgt	gtaacaggcg	420
gaccacctgt	tgggactgcg	agattctcaa	ggggaaggac	tgggtctcat	ttctcccac	480
tcagcgctta	gcaggatgac	ctggtataga	gcagggaact	gggaaatgtg	ggtcaggggg	540
tcagacactc	cagttgggtc	ttttatataa	attaaatggc	aaaagggtcc	atacccttct	600
ccttctttcc	taccctccac	tttatctgca	aaatgggaat	gatgataaca	cccacttcat	660
agaatgggtc	tgaagatcaa	atgagagaat	aaaagtcaag	cacttagcct	ctggtgcaca	720
attagttatta	aataagtata	cctattccct	cttttctttt	tttaaaaaata	atattaccaa	780
atgtccagct	tatacacatt	tacaagactt	agctagtggg	ctatgttaga	gctactaaaa	840
gatctttgac	aagctaaaa	taagatgcaa	tgaatgaggt	gtaacgaaca	agagagtttt	900
aagttcagaa	atggttacag	aagtataaga	cagctgtgtg	ggtgtttttt	gatttttggt	960
ttctggttta	caatctcgtc	attcaacaaa	gatgggagtt	ttatagaact	aaaagcacca	1020
tgtaaagctac	taaaaaaca	aacaaaaaag	gctcatcatt	tctcagctctg	aattgacaaa	1080
aatgccaatg	caaataaaaa	tgattacttt	ttatttttaa	aaaaaaaaaa	aaaaaaaaaa	1140
aaaaaaaaaa						1150

<210> 70

<211> 1398

<212> DNA

<213> Homo sapiens

<400> 70

gggagatagg	aatggaattg	atggctttat	tcttcagaac	taccaccgta	gcagccatgg	60
caagcagggg	cgccttgcg	ttgttcttga	ggaaaatatt	aagtgaggct	aaattcaaac	120
tgagcctgac	tcccaaaccc	ccacaaccct	tttatatata	tatggcatat	tacagtgaga	180
atttcttttt	aaagttttag	gtctttgtct	ttttcttgat	aaaattatag	ttataatagt	240
tgtcataagt	taatcaatcg	tatgttttaa	gtgcccttag	tgcaaaatth	gatgcccttg	300
gatactgttg	atttattaat	atgaaatata	cctttgttaa	tttttaattt	tatggatagg	360
tgtatatgtt	tatggggtgt	gagatacaac	ttactctatt	tacgttactt	accgcagaac	420
cacatatttc	ttccaaatgt	ccatgcccta	tcttagttac	aaaaacacaa	tcatttgcca	480
agtgaattat	cggtgattat	tagaaatcat	ttatactgtg	gttctgtgtg	cattttaaact	540
ataatatact	tagaaattat	tcttgctttg	gtaaaatact	ttccttctta	ttcaagagga	600
ttttttgggt	catgctgtct	tgttatatac	tcttttaagt	ttattatgat	tgtgtgtgta	660

40

tatgtttgtt	ttttcttttt	ccttctgtct	gaattctgtt	gcactgagca	atgttgtaat	720
atttttat	ttaaataaag	taatatataa	aattactgga	aatatgtaac	catcagatta	780
ttatctccta	atgataaaca	gaatttggtta	attaagctaa	acctagaatt	gtagacaatt	840
atttttacat	tgcactctaca	ttaaaatgct	atctcaaaca	cacatacttg	gttgtgtaat	900
atttatctac	tcattaagta	gaaagagtaa	attaaaaaatt	gcttttgat	tattgatgag	960
ggtggattat	actttagaac	actttattca	aacagttctt	ccacatatct	cccttttgac	1020
ttgactgagc	aactctcttt	ctgtgcttcg	gtttgggtctc	taagtcagag	ttaatatattc	1080
ttgctctatc	tagcatatag	aagcattgtg	ggctgggtgc	agtagctcac	acctgtaatc	1140
ctagcacttt	gggcagattg	cccaagctta	ggagtttgag	atcggcctgg	gtaacatggc	1200
gaaatcccg	ctctactaaa	aacacaaaaa	aattagctgg	gtatgggtggc	gcacgcctgt	1260
aattccagct	acttggaag	ctgaggcgca	agaattgctg	gaacctggga	ggcggaggtt	1320
gcagtgagcc	gagatttcgc	cttgcactcc	agcctggcga	gattctgtct	ccaaaaaaa	1380
aaaaaaaaa	aactcgt					1398

<210> 71

<211> 1557

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1541)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1549)

<223> n equals a,t,g, or c

<400> 71

gcaaaggtga	agctgggtttt	catgggtctcc	tgaggggcccc	tggccccctgg	gagatggggtc	60
acactccctg	aatgctgtgc	tggtgggttc	cctggaggat	tcttgctgca	ggccagggtcc	120
cgtattctcc	acactcacca	caagtggctg	gggtgtgactt	gacacgggtg	gaaagtggag	180
gggcgcgagc	actcagtatc	cagcgcgagc	cattgggtgtg	cctagaaaaat	tactacaaag	240
atttcacat	ctataaccca	aacctcctaa	cagcctccaa	attccgagca	gccaaagcata	300
tggccgggct	gaaagtctac	aatgtagatg	gccccagtaa	caatgccact	ggccagtccc	360
gggccatgat	tgctgcagct	gctcggcgca	gggactcaag	ccacaacgag	ttgtattatg	420
aagaggccga	acatgaacgg	cgagtaaaga	agcggaaagc	aaggctgggtg	gttgacgtgg	480
aagaggcctt	catccacatt	cagcgtctcc	aggctgagga	gcagcagaaa	gccccagggg	540
agggtgatga	ccctagggag	gcggcccagg	ccattttccc	ctccatggcc	agggtctctcc	600
agaagtacct	gcgcacacc	cggcagcaga	actaccacag	catggagagc	atcctgcagc	660
acctggcctt	ctgcacacc	aacggcatga	cccccaaggc	cttcctagaa	cgttacctca	720
gtgcggggccc	caccctgcaa	tatgacaagg	accgctggct	ctctacacag	tggaggcttg	780
tcagtgatga	ggctgtgact	aatggattac	gggatggaat	tgtgttcgtc	cttaagtgtc	840
tggacttcag	cctcgtagtc	aatgtgaaga	aaattccatt	catcatactc	tctgaagagt	900
tcatagaccc	caaatctcac	aaatttgctc	ttcgtttaca	gtctgagaca	tccgttttaa	960
agttctatat	ttgtggcttt	attaaaaaaa	aaaraaaaaat	atatagagag	atatatatct	1020
atgccagagg	gggtgtctttt	ttaaaaaatc	ttcttcattg	ctgactgaaa	ctggcagatg	1080
attgaccagt	atcctttgac	catctgcact	ttatttgga	ggaagcaggg	gctgtccacc	1140
ctgaaaaaga	gtgactgatg	acatctgact	tttgtcgtatg	ggacttctca	agaagccatt	1200
ccttgagct	tctgttacag	ctgtaaacca	aagtggagct	gggtgttctt	gggagcctcg	1260
ccttacaact	agttcctgcc	tttcgtccag	taccaagtcc	cccgttgctt	ctggctcagcc	1320
cacttgtaga	cttccagggg	acacatcttt	attctgtttc	aggaaaccag	tcacracacg	1380
tccacatatg	tatttggtga	tgtaaatgcc	agtatcacat	cacccatgaa	agtcgtgggc	1440
agttcargag	atacctgsct	tcgtctttgk	tctttgttgc	cttaggttct	tcagagaaag	1500
atcmcaacaa	aaaatgtacm	ctgtcgttta	cagctawaag	ngatttgant	tgttttt	1557

<210> 72

<211> 1163

<212> DNA

<213> Homo sapiens

<400> 72

ggcacgagct	ggctgcaggg	tctctgggga	gagaaggggc	ctcggttca	caggatgggg	60
ctgccagtgt	cctgggcccc	tcttgccttc	tgggttctag	ggtgctgcgc	cctgctcctc	120
tcgctgtggg	cgctgtgcac	agcctgccgc	aggcccgagg	acgctgtagc	ccccaggaag	180
agggcgcgga	ggcagcgggc	gaggctgcag	ggcagtgcga	cggcggcgga	agcgtcccta	240
ctgaggcgga	cccacctctg	ctccctcagc	aagtcggaca	ccagactgca	cgagctgcac	300
cggggcccg	gcagcagcag	ggccctgcgg	cctgccagca	tggatctcct	gcgcccacac	360
tggctggaag	tgtccagga	catcacccga	cgcgagcgag	ccccctctgc	cttcccacac	420
caggagctgc	ccggggctct	gcgggcagct	gcagccaccg	caggggtgcgc	tggcctcgag	480
gccacctatt	ccaacgtggg	gctggcgggc	cttcccgggg	tcagcctggc	ggccagccct	540
gtggtggccg	agtatgcccg	cgtccagaag	cgcaaaggga	cccacgcag	tccccaagag	600
ccacagcagg	ggaagactga	ggtgaccccg	gccgctcagg	tggacgtcct	gtactccagg	660
gtctgcaagc	ctaaaaggag	ggacccagga	cccaccacag	acccgctgga	ccccaaaggg	720
cagggagcga	ttctggccct	ggcgggtgac	ctggcctacc	agaccctccc	gctcagggcc	780
ctggatgtgg	acagcgcccc	cctggaaaac	gtgtatgaga	gcatccggga	gctgggggag	840
cctgctggca	ggagcagcac	gtgcggggct	gggacgcccc	ctgcttcag	ctgccccagc	900
ctagggaggg	gctggagacc	cctccctgcc	tccctgccct	gaacactcaa	ggacctgtgc	960
tccttcctcc	agagtggagc	ccgtcccccg	ccccgccccg	cctcacagct	gacagcgcca	1020
gtcccaggtc	ccggggccgc	cagcccgtag	ggtccgtgag	gtcctggccg	ctctgacagc	1080
cgcggcctcc	ccgggtcca	gagaaggccc	gcgtctaaat	aaagcgccag	cgcaggatga	1140
aagcgaaaaa	aaaaaaaaaa	aaa				1163

<210> 73

<211> 1486

<212> DNA

<213> Homo sapiens

<400> 73

cggcacgagc	cagggctgag	gtaggagggg	gtctgtccct	cgacgcctcc	tgcgacgcca	60
gcccctgagc	gatgatgcga	acgtgcgtct	tactctccgc	ggtgctctgg	tgcctcacag	120
gagtccaatg	ccgcggtttt	accttattca	ataagaaggg	cttcatttat	ggcaagacag	180
gacagccaga	caaaatatat	gtagagttag	atcaaaatag	tccagtcctt	atctgtatgg	240
attttaagct	ttctaaaaaa	gaaatagtgg	acccaccta	cttatggatt	gggcctaattg	300
aaaagacggt	aacaggaaat	aatagaataa	atataactga	aactggacag	ctgatggtga	360
aagatttttt	ggagcctttg	tctggacttt	acacatgtac	tctttcttat	aagactgtta	420
aagcagaaac	tcaagaagaa	aaaacagtca	aaaagagata	tgactttatg	gtctttgcct	480
atcgggaacc	tgattattca	tatcagatgg	ctgtacgttt	taccacaagg	tcttgtatag	540
ggagatacaa	tgatgtattc	tttagagtgc	tgaagaaaat	cttggatatt	ctaatttctg	600
atttgtcatg	ccatgtcata	gagccatcat	ataaatgcc	ttctgttgaa	attccagaac	660
atggcctcat	acatgagcta	tttatagcat	ttcaagttaa	tccttttgcg	ccggggtgga	720
aaggtgcttg	caatggatct	gttgactgtg	aagataccac	taatcataat	atcctccagg	780
caagagatcg	aatagaagac	ttttttcgga	gccaaagcata	tattttctac	cataacttta	840
ataaaactct	accagcaatg	cattttgtgg	accacagttt	gcaagtagta	cgtctggata	900
gctgtcgacc	aggccttgga	aaaaatgaac	gtctacacag	taattgtgct	agctgttggtg	960
tggttttag	tcctgcgact	tttagtcctg	atgttaatgt	aacttgctcag	acctgcgttt	1020
ccgtccttac	ctatggagct	aaatccttgc	cacaaacttc	aaacaaaaat	cagcaatatg	1080
aagattagag	gtgaaagcat	tgttacttac	ttgtggaagt	cggggacata	agatgatcctt	1140
cacatcccag	agcatcatag	atagttccat	taagtaaaat	cagtaagacc	aaaccactgg	1200
gaaaacatgc	attttggaag	ctttaaaata	aaatgggtta	acatggcatt	tctaagaagg	1260
catttaatcc	ggtatctcta	gtgtacaaag	gaaacttgaa	gttttcatgg	attattttta	1320
atgaaatgtt	ttattgttta	caaacggtat	gttgctgtgt	acctaagggg	ttagtaagggt	1380
caagaagggt	ttcaaagttt	aataaaataa	aaattgtata	ctctccaaaa	aaaaaaaaaa	1440
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaa		1486

<210> 74

<211> 1553

<212> DNA

<213> Homo sapiens

<400> 74

ggcacgaggg	gatgcagcag	agaggagcag	ctggaagccg	tggctgcgct	ctcttccctc	60
tgctgggcgt	cctgttcttc	cagtgagtg	ctccggctgg	ctacgctcca	cttccggccg	120
gggggttagg	gaagatggtc	gcttttccgg	ttccgggtag	gggggtctcc	agaaagcccc	180
cgcatagctc	tgggaaggaa	ggaggaggg	agcgggacgt	tgggacgatg	tcatcaccac	240
cccgttgagg	atagttggta	ttttgtcagt	cctttcctga	gccttgggag	ctttttcacc	300
ttgcgtgacc	tgtcattgta	tttggtgata	catacttgat	ggaacacaaa	tagctctggc	360
tttgggggca	cccgggtgta	ttttgtggtg	ttattatcag	tggttgaccg	ctgtcgtttg	420
ggcgctgtat	accctgatta	gaaagaaaac	agttctagca	ttcagtagtt	tgccccagta	480
gttttgaggaa	acaaaatcat	gacagttggg	aattttat	tttgtagtga	ttattgtagt	540
ctataaaatg	aaatatttgc	taatcacaga	aaggatgtga	tgtgattcgt	tatagaaaaa	600
catatcgatg	aaaaagtact	gctttagcat	ttcataactt	tttttaatac	aagaaagagc	660
atataaattc	tggtaaaatt	tattaagtaa	tatcagaagt	agggtattca	cactgacatt	720
cacatcttaa	aattaggtga	gttttgcttt	catataattt	tcagattgca	taaatttcaa	780
tccagcctaa	ttaccttcat	ctcaagggtac	aattatttcc	tggattttgt	gtatgtgtgt	840
cttctctggga	aataaaatat	tagtatatct	agtattagag	tgaattgcta	aattttataaa	900
atttctcttga	gggtaatggg	gatttctgaa	agagataatt	aatattatat	aaatataaag	960
actttaatat	aaagacttca	gatgtcctaa	atcttagcag	ttaaggctgg	aaagttttta	1020
actgttgact	aatattcatg	ttgccagcac	acccccgatc	cctcccagct	gcctgtgtca	1080
tcataagtct	ccagttttta	agctaaatta	ctttgcagta	tgatatttag	ctggcctact	1140
gtcactccta	ataaaccag	tcctctctct	ctctctttct	cacacacaca	cacacacaca	1200
cacacacaca	cacacacaca	caacagagaa	atattcaaat	attgtcttat	tgagaaaaaa	1260
aatctgga	tggtgttttc	ttaaatctaa	tgatcatgtt	ctttgaagag	catctgtcaa	1320
gtccagcaga	ccatataatt	caatcaaaaa	tttaaaagtt	atatttacaat	ctctatgtac	1380
ttgtggaata	gattttttgc	cttttaaaaa	ataactgggt	gcaggggtgt	gtggttcttg	1440
cctgtaatcc	cagcactttg	ggagaccgag	gcagctggat	cacctgaggt	caggagttcg	1500
agaccaggct	gaccaacata	gtgaaactcc	gtctctacta	aaaaaaaaaa	aaa	1553

<210> 75

<211> 1650

<212> DNA

<213> Homo sapiens

<400> 75

ggaacctcat	caacgctgac	ttctgctgtg	cctctgtctg	cgtggccttt	ggggcagttc	60
tgggtaaagt	cagccccatt	cagctgctca	tcatgacttt	cttccaagt	accctcttcg	120
ctgtgaatga	gttcattctc	cttaacctgc	taaagtgaa	ggatgcagga	ggctccatga	180
ccatccacac	atttggegcc	tactttgggc	tcacagtga	ccggatcctc	taccgacgca	240
acctagagca	gagcaaggag	agacagaatt	ctgtgtacca	gtcggacctc	tttgccatga	300
ttggcaccct	cttctgtgtg	atgtactggc	ccagcttcaa	ctcagccata	tcctaccatg	360
gggacagcca	gcaccgagcc	gccatcaaca	cctactgctc	cttggcagcc	tgcgtgctta	420
cctcggtggc	aatatccagt	gccctgcaaa	gaagggcaag	ctggacatgg	tgcacatcca	480
gaatgccacg	ctcgcaggag	gggtggccgt	gggtaccgct	gctgagatga	tgctcatgcc	540
ttacgggtgc	ctcatcatcg	gcttcgtctg	cggcatcctc	tccaccctgg	gttttgtata	600
cctgacccca	ttcctggagt	cccggtgca	catccaggac	acatgtggca	tttaacaact	660
gcatggcatt	cctggcatca	taggcggcat	cgtgggtgct	gtgacagcgg	cctccgccag	720
ccttgaagtc	tatggaaaag	aagggtctgt	ccattccttt	gactttcaag	gtttcaacgg	780
ggactggacc	gcaagaacac	agggaaagtt	ccagatttat	ggtctcttgg	tgaccctggc	840
catggccctg	atgggtggca	tcattgtggg	gctcattttg	agattaccat	tctggggaca	900
accttcagat	gagaactgct	ttgaggatgc	ggtytactgg	gagatgcctg	aagggaacag	960
cactgtytac	atccctgagg	accccacctt	caagccctca	ggaccctcag	taccctcagt	1020
acccatgggt	tccccactac	ccatggcttc	ctcggtagcc	ttggtagcct	aggctcccag	1080
ggcaggtgag	gagcaggctc	cacagactgt	cctggggccc	agaggagctg	gtgctgacct	1140
agctagggat	gcaagagtga	gcaagcagca	ccccacctg	ctggcttggc	ctcaaggtgc	1200
ctccacccct	gccctcccct	tcatcccagg	gggtctgmct	gagaatggag	aaggagaagc	1260
tacaaagtgg	gsatccaagc	cgggttctgg	ctgcagaagt	tctgcctctg	cctggggctc	1320

43

tgccacatt	ggagaaaaac	aggctcaaag	tggggctggg	acctgggtgg	tgaacctgag	1380
ctctcccagg	agacaactta	gctgccagtc	accacctatg	aggctcttct	accccgaggc	1440
tgcacctcgg	ccagcatctc	ctatgctccc	tgggtccccc	agacctctyt	gtgtgtgtg	1500
cgtggcagcc	tccaggaata	aacattcttg	ttgtcctttg	taaaatgggtg	tgaatgctcc	1560
aatggggcca	gtttgaggga	gaaaaggacc	caagagacct	gcttctgccc	cagcccttac	1620
cttcatccaa	gggtaccaac	cacactgcga				1650

<210> 76

<211> 2150

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (874)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1198)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1201)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1266)

<223> n equals a,t,g, or c

<400> 76

ccacgcgtcc	ggacccgagc	tccagtagtt	ccgcccgtg	gtcatcgcg	cctttccct	60
gcccgtgtcc	tgtcgcggt	ccccgccatg	ctgtctctag	actttttgga	cgatgtgcgg	120
cggatgaaca	agcggcaggt	gagcttgtcc	gtcctctttt	tctcctgggt	cttcttgtcc	180
cttcgaggct	gctgctgcgg	ggccggcg	accccgagg	tctggtgtga	gggtctgagc	240
tgggtctgata	cccggggtcat	tcgctttctt	tggagactgt	ggccagaggc	cgccttgtcc	300
gcctcattat	ttttaacccc	gaactgattc	aggccctacc	tggggcgggg	cgggaagcgg	360
tgtcttcacg	ttccattcct	cccactgagg	caggggagca	aatggaaacc	gtacgcgctt	420
gaagtgggag	ttgggggtgct	tattgtttta	gtcattttta	tgcggcggac	tcttgatttc	480
tccagtcgga	gcgactccag	gtggtttcgg	gagagacgag	gtttagccgg	tttctggggc	540
gctcaggaag	gcgattggag	gccccacaaa	aaccgttttg	ctgctttcag	ctccttgcaa	600
cccttttagta	gagctgaacc	gtagcgggct	gcaccgactt	tgacttggac	cactctgggc	660
tccgagttgg	aacagttaca	ctacttgccc	ttgcgtccgc	ttagcactaa	ggcggcagcc	720
ctcggaatct	atgggttttac	agtccaatat	cagtgccacg	gggatctgga	aatgtaggtc	780
tcctgatttt	gtccttacac	tttactttga	tcttctagat	cgatgcca	atagtactga	840
gaatattgtt	gtaattattt	agtccttaga	aaangttggt	ctgttttatc	ttttgcgcct	900
agtgtgtctg	tagagcctag	ttttgctgca	tcggactttt	tttttgtttt	aaacagtatt	960
ttactgttat	gattatcctg	atgtcaccat	taaggatttt	ttttttcctt	ggacttgcac	1020
tttttgtagt	tataactgcc	acttagggaa	gtagatacac	aacctttcct	tactcccctt	1080
caggccttag	ctagctcagt	gtcaattctg	tcagtcagaa	ttgagcattc	tataaaaatt	1140
gcgcaaacgt	tactttatgt	ctttatgaca	acacttcaaa	tttttacttg	tatagtntg	1200
ncttttttta	atccatattt	ggatttctag	atgccacaga	tatttctctg	aggaaagtat	1260
ttatnttgag	tctgatattt	attgactcta	tgtaggtcc	aatgagagaa	atgcaaagat	1320
agttaagaaa	gactcggcct	tcaaggagcc	taaatgtgta	gaaaaggact	aaggcaaac	1380
aataactttt	ttgagctcct	gccatgtgtg	aagcacttta	tacacctgta	aggtaggtaa	1440
cgttggttct	attaaacatg	aagaaaatga	gactttgtga	gaagcaatac	agtatagaag	1500
ttaagaatat	ggactctaaa	gctagatttc	agaggtttga	agtagctctg	ctacttactg	1560

44

gctgtgtgac	tttgagcaga	ttacttaacc	tgtctgtgcc	tatgtttact	tttattgttg	1620
taaaaagata	tgcaacataa	aatattccat	ttcaaccgtt	tttacgtgta	tacttcactg	1680
acattagttg	cattcactat	gttgtgcaaa	cgtagggctg	ctatgaagat	taaatgagtt	1740
aattcatata	aagccctcag	aagagtgtct	ggcacatggt	gagtattggc	tgtactgtgg	1800
tcgatgtcat	tgtagagag	ctttagtgtat	ttgcttaaga	cagaaggtag	actgggtgctg	1860
ggtaggtcac	gcctgtaatc	ccagcacttt	gggaggctga	ggcaggcgga	tcacaatgtc	1920
aagagattga	gaccatcctg	gccagcatgg	tgaggccccc	tctctactaa	aaatacagat	1980
actagctggg	cctgttggtg	cacgcctgta	gttccagcta	ctcaggaggc	tgaggcggtg	2040
gaatcgctg	ggaggtggag	attgcagtga	gctgagatcg	tgccaccgca	ctccagcctg	2100
gtgacagagt	gagactccgt	ctcaaaaaaa	aaaaaaaaaa	aaaaaaaaaa		2150

<210> 77

<211> 1592

<212> DNA

<213> Homo sapiens

<400> 77

cggatttgtt	gagtgaatga	gaatagtcag	tcaagactct	ggattgttga	atttttaggtg	60
ctgctcattg	tccctctttt	atcttcagaa	atctcgatga	aaatatacac	atagaagaca	120
aaaccgaaga	aatattttat	tatttgttgg	ggcttctcta	atttttgtat	agcttttagta	180
agctgaatta	gatggcatgg	aaccgagaag	tttccttcta	cctgaattgg	gtgggagagt	240
gtcacacatt	cctcttggcc	tcactctggt	ttttgcctgc	tttcttatgg	ttagggagac	300
tgcaggaggt	tttagcttca	gagcaggaga	cttagaagaa	atctcaagaa	agagaacaaa	360
tgtattaggg	tctcttagag	ggacagagct	aataggatat	atataatcct	attatatata	420
tacacagaca	cacacacaca	tatatatata	tacacatata	tacatatata	tacacacata	480
catgtatata	catatatata	cacacatata	catgtatata	catgtatata	cacacatata	540
catgtatata	catgtatata	cacacatata	catatatata	catatatata	cacacacata	600
catgtatata	catatatata	tacacacata	tatgagttta	ttaagtatta	atttacatga	660
tcacaaggct	ccataataga	gtgtctgcag	ctgagggcaa	ggagagccag	tccaagtcct	720
aaaattgaag	aaattggagt	ccgacatttg	agggctggaa	gcattcagca	caggagaaaag	780
atgtaggctg	agaggctagt	cctgtcttgc	cttttcacat	ttttctgctt	gccttatgtt	840
cactggaaga	tgattaaatt	atgcacacta	gattaagtgc	agatgtgcct	tccccagccc	900
actgactcaa	atgttaattct	cttttggcaa	aaccacaacag	acacaccag	gattaatagt	960
ttgtgtgctt	cagtccaatg	aaattgacat	tcagtattaa	ccatcacaa	aaacgtgcat	1020
ttacattgca	tagtgaatgt	ccttgaactt	gttcagttcca	gtttcagggtg	tcttatttct	1080
tgggggtaga	ataagaaaag	cagagagaag	tcttcttttc	ccttggagca	ctggttctca	1140
gatactttta	ttttctagcc	tcttatgggg	taataataat	taaaatggag	aaaaataaaa	1200
taatagtaaa	atttgaaaca	atgcaaaaaa	tgtaaaaacg	aaattactga	aaaaagggcc	1260
cccattcgcc	agaagtaggc	aggtggctcc	cagaggtggg	cagagcttag	gctgaaggag	1320
gcgggctgag	ctgcacctgg	ggtggagagg	gcctgagtc	actcctagtt	aaggtctgtc	1380
aaacccatct	gcaccgctca	tctgccgcag	gggaggctgt	gagggggcag	aagcagctcc	1440
tctccctgag	ggctagccat	ctggagggag	acctacagac	ctaattgtta	tgctgcatgc	1500
tggaatagaa	ggaatcgtgt	ggtgctgagg	gaacctgtcc	cctaaccag	accagaggag	1560
gtccggaatt	cgatatcaag	cttatcgata	cc			1592

<210> 78

<211> 1579

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1529)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1556)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1569)

<223> n equals a,t,g, or c

<400> 78

ggcagagggg	acccacgcgg	aggaaggaag	agacgcaggc	aggctgcggt	tacccaagcg	60
gccacccggg	cctcagggac	cccttccccg	agagacggca	ccatgaccca	gggaaagctc	120
tccgtggcta	acaaggcccc	tgggaccgag	gggcagcagc	agggtgcatg	cgagaagaag	180
gaggctccag	cagtgccctc	agccccaccc	tcctatgagg	aagccacctc	tggggagggg	240
atgaaggcag	gggccttccc	cccagcccc	acagcgggtc	ctctccaccc	tagctgggcc	300
tatgtggacc	ccagcagcag	ctccagctat	gacaacggtt	tccccaccgg	agaccatgag	360
ctcttcacca	ctttcagctg	ggatgaccag	aaagtctcgt	gagtctttgt	cagaaaggtc	420
tacaccatcc	tgtgtattca	gctgctgggt	accttggtct	tctgtggctct	ctttactttc	480
tgtgaccctg	tcaaggacta	tgtccaggcc	aaccaggtct	ggtactgggc	atcctatgct	540
gtgttctttg	caacctacct	gacctgggt	tgtgttctct	gacctaggag	gcatttcccc	600
tgggaacctg	attctcctga	ccgtctttac	cctgtccatg	gcctacctca	ctgggatgct	660
gtccagctac	tacaacacca	cctccgtgct	gctgtgcctg	ggcatcacgg	cccttgtctg	720
cctctcagtc	accgtcttca	gcttccagac	caagtctcg	ttcacctcct	gccagggcgt	780
gctcttcgtg	cttctcatga	ctcttttctt	cagcggactc	atcctggcca	tcctcctacc	840
cttccaatat	gtgccctggc	tccatgcagt	ttatgcagca	ctgggagcgg	gtgtatttac	900
attgttcctg	gcacttgaca	cccagttgct	gatgggtaac	cgacgccact	cgctgagccc	960
tgaggagtat	atttttggag	ccctcaacat	ttacctagac	atcatctata	tcttcacctt	1020
cttcctgcag	ctttttggca	ctaaccgaga	atgaggagcc	ctccctgccc	caccgtcttc	1080
cagagaatgc	gcccctcctg	gttccctgtc	cctccctgtc	gctcctgcga	gaccagatat	1140
aaaactagct	gccaacccag	cctgtggcca	ggtcactgtc	taccccagcc	cagcccagcc	1200
ctctgccgct	tgtacatacg	ccatggggac	cctgaggaac	tgaggccacg	tcaatccctg	1260
tgccgcccga	ttcgcccggt	acatcttcca	aactgggacg	gtcaaggctg	aaggctcctc	1320
tgggttttag	ggtccaaggg	acaaggagga	gaagcctagc	aggatttcag	atgcaggaga	1380
gagacccagg	aagcccggca	gagcctgagc	cccaytgcaa	tttyctyctag	ggstgcacaw	1440
tcatgtggcy	ttagggcama	ytgtctgca	tccagtctgt	gttyctyctgt	ctttctcatc	1500
caggtcaggc	attgacattt	gtaagaaang	gggtaaggga	cacagctggg	caagtngatt	1560
ggttggcang	attgctgtc					1579

<210> 79

<211> 1396

<212> DNA

<213> Homo sapiens

<400> 79

ggcacgagaa	aaatatgaag	tgcacagctg	tgtttgcacc	atcagcatgg	ccaaacaccc	60
tttctcttct	cgtttctctc	cacacagtga	tgtgcattaa	ctggcacctg	gtttctgcat	120
cacacatgca	tattgggaga	attgttattc	tagaaggggg	tggaaatgtg	cacacctcag	180
ataattgctc	tgagaaatga	gtatgaggtc	tttttaaaaa	aactttacat	ttaaaccacac	240
aaaccattta	aaaaaaccac	tcctccaaaa	cagacacaca	agtgaacaaa	accaaatacga	300
ctgtccatcc	atttaatttt	gaactatctg	caatatcacc	ttaaagctct	aagctcctaa	360
tgatttttgg	cattacaaat	ccaaaataac	attatttcaa	aatgaagttt	tactatgatt	420
tttcccatth	gggaaacagt	atatccaata	caacaagatg	catgtgtacc	tccattaaaa	480
ttttaacaat	taggaaaatt	agaatttata	tatgatagat	taaaatatga	ttatttccta	540
aaagtthtct	ttagttttac	aggtttgctt	ctggattgct	ttcaatagtt	attagtgaat	600
ttaaaatttg	attcagtatg	aaggtaatgt	acacccatct	tcttcagaac	acgtatttta	660
aaatgtgaca	tcatctgttc	tgagtctaa	cactgaagaa	tgaagtgtcc	acattcaccg	720
tattttatga	cagacacatt	ctcattccat	tgacatttaa	gtgatgtaat	gaactaaaac	780
tgccacaaag	agaggagggt	tgggaggctg	agctctggac	agctctgtca	ctgggcaagt	840
aacatgggtg	ttagaaaatc	tctgtctcct	tctgtgccat	tggctctatat	ctctgttttg	900
gacaaaaaac	caaacaccgc	atgttctcac	tcatagggtg	gaattgaaca	atgagaacac	960
ttgcacacag	gaaggggaac	atcacacacc	ggggcctgtt	gtgggatggg	gagagcgggg	1020
aggaatagca	ttaggagata	tacctaatgt	aaatgacgag	ttaacgggtg	cagcacacca	1080

46

acatggcaca	tgtatacata	tgtaacaaac	ctgcacattg	tgcacatgta	ccctagaact	1140
taaagtataa	taaaataaaa	ataaaaatta	agttaaagtaa	aataaaatct	ctgctcctgt	1200
ttgctgtccc	cactgataaa	atggtagagc	agatcactcc	taagtcagag	cttgtgtaat	1260
agctgcgtag	gagtgtgaaa	ccaggctact	tgggagccag	tgctggctcc	agaactttcc	1320
agctgcgtgt	ccttgggkca	ggtgggtttac	actctgttgc	ctcagtttct	ccagcaataa	1380
aaaaaaaaaa	aaaaaa					1396

<210> 80

<211> 1230

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1223)

<223> n equals a,t,g, or c

<400> 80

cagcaccatc	gcctacctga	cctcccagct	gcacgccgcc	aagaagaagc	tcattgagctc	60
cagcgggacc	tcagatgcca	gcccgtcagg	gagccccgtg	ctggccagct	acaagccagc	120
gcccccaaaa	gacaagctac	ccgaaacgcc	tcgccgccgc	atgaaaaaga	gcctctcagc	180
ccccttgac	ccggaatttg	aagaggtcta	cagattcggg	gcagagagca	ggaaactcct	240
tttgcgggaa	ccagtggatg	ctatgcccg	ccccaccca	tttytgctgg	ctaggaggatc	300
cgccgaggtc	cacctcatca	aagagaggcc	cctcgtcatc	ccccccatcg	cctccgaccg	360
aagcggcgag	cagcacagcc	cggcccgcga	aaagccgcac	aaggcccacg	tcggggtggc	420
acatcgatc	caccacgcca	ccccgccgca	gccagcccga	ggtgaagacc	ctggcggctc	480
accaggtgaa	cggaggcaag	gtgggtgagga	agcactcagg	gacggacaga	actgtgtgaa	540
gcccgcctg	ccccaccccg	cgtgttccat	gcactgtgag	caccactggg	aaatctcagc	600
cacacctttt	ctgtttaatc	ccatgcatgc	caaacacttt	tcacacctac	cgaccatttc	660
tccttctgct	tctcttggcc	tcttcttcac	acaaaaatat	gatcgtgtcc	ctgccgcaga	720
atatgtattt	cctaattgct	gtggccaaac	gcctgtgtgc	cgaatcgctt	gcttctgata	780
ccgctccgtg	taacctaaag	gcgctgcagg	caaagcccag	gccacggctg	cgctcactact	840
gatgttcacg	atgccacaca	gtcacacacc	taattcattc	tcaagtcgga	gcaacacata	900
ccaaccttga	ccttatccct	aagctccagg	gcagcctggc	cgagcagccc	ctgctccctc	960
ctggagaccc	ttgtcacctc	ccgagctcct	cctggagacc	cctgtcacct	cctgaccaac	1020
ctttcccagg	gcggcaccga	tcaccgagca	gccgtgcgtg	tatctcaagg	aactaaataa	1080
gatgacgcta	ctcctcatag	caccacaacc	tgaatgtgtg	ttcatatttt	tttgttagtt	1140
ttatccaaaa	tgtttaagat	cccaacaaac	tttattttct	aaacctgcaa	aaaaaaaaaa	1200
aaaaaaaaaa	aaaaaaagg	ggnccctttt				1230

<210> 81

<211> 1139

<212> DNA

<213> Homo sapiens

<400> 81

acgcgtgggt	ccggacgcgt	gggaggagc	gtgggagcaa	gcccaggcgg	cggtggaaaag	60
gctggaggac	acacctaaac	atgtggaatc	ccaatgccgg	gcagccaggg	ccaaatccat	120
atcccccaaa	tattgggtgc	cctggagggt	ccaatcctgc	ccaccacca	cctattaatc	180
caccctttcc	cccaggcccc	tgtcctcctc	ccccaggagc	tccccatggc	aatccagctt	240
tccccccagg	tgggccccct	catcctgtgc	cacagccagg	gtatccagga	tgccaaccgt	300
tgggtcccta	ccctcctcca	taccacccgc	ctgcccctgg	aatccctcct	gtgaatccct	360
tggctccttg	catggttggg	ccagcagtga	tagtagacaa	gaagatgcag	aagaaaatga	420
agaaagctca	taaaaagatg	cacaagcacc	aaaagcacca	caagtaccac	aagcatggca	480
agcattcctc	ctcttccctc	tcctcttcca	gcagtgatcc	tgactgaata	caggccctgg	540
acccttccct	caagtctcac	cagttctgct	ctcccatcaa	gcttcagatg	ccatgttgta	600
ctgggggaat	gtagcccttg	tgctccccc	cccctaccts	cacctgagcc	tcaccctgct	660
gttgagccct	gagtggctag	gggaaatggg	aagaggattg	ccatggcctg	gccatcttgt	720
tgctgcttgg	ttagatcata	tagctaata	attaggcagg	ggagctatct	tttgaagatg	780

47

atgaactaaa	tggtgaagac	aagtttgaga	tctgtaaaa	gtgatttttt	acttccactt	840
ataatacttg	tgattgggga	ggtttggtga	aattcaatta	tgatgaaaaa	cctatctttt	900
ttgtaatggt	ggcatacttg	gggaatttag	tggcaaatac	attccccagc	aggccttttg	960
ttggttgac	taactgcaag	gttgctggga	agtagagtc	atttggttga	tgagctttga	1020
ctgcggtttt	ggaaccttac	ctctcctcct	tagcccaata	tgctgtcttg	ggctctattc	1080
aaataaagtt	atttctcctg	gtcwmaaaaa	aacggcacga	gcggcacgag	ctacgtggg	1139

<210> 82

<211> 1409

<212> DNA

<213> Homo sapiens

<400> 82

ggcacgagga	acctcccgcg	ggctttggac	tgaggtcctt	gtggcgctcg	tctcctcccc	60
atgaagtggg	agcgaggctc	cccaatggtg	cttttggtt	tagtgtacga	tgtttgctgt	120
gcttcccgc	gtggagggca	gagccacccc	acatcaggat	cggacgtgct	acccctcccc	180
gtccccgccc	tggcccagcc	agcccagccc	tcgaggctcg	atgcctgtgc	caaggccagg	240
ggcagccaga	ggcagctgg	atggccacgt	gcagggtcaa	ggctgggccc	tgagtgggg	300
cggggccgcca	gccccagcag	tttacagacg	catggctctt	cctcccagag	cagccggcag	360
ctacctggac	cggaaatgtc	ctcatccctt	ccctggggcc	aggctctgcc	ctggccttcc	420
tctgtgaacc	cctcctttct	ttgtgctgtc	tcgggactcc	tgaccgtggt	gtgcgtgtgt	480
gcccgtctgt	gactttctac	tcaccaaggg	ttgaagaaag	gaaacgggga	aaatcaaaag	540
gggttcaaac	cccacctcag	taggtggagg	ggagcgcttg	ccattgggtg	tatttttgtt	600
ctgagttttc	ggtgcctgt	tcctaactac	tccatcccat	gacctcgcca	cacctactgg	660
ggcatctggc	tggtgcctgc	tgccatggcc	agccccact	ctcaccctgc	acagggggtc	720
ttgcagcccc	caggcccaca	gcctcgttgg	gaggacaggg	tggccctggg	gacaagaggg	780
aggagcccag	gggcttacct	cactgagagt	gtcccccagc	aggcatccac	taccccaggg	840
ccccccacat	gtcatggcaa	ggttggtagt	gaatgggcct	ggttgggagc	agccccaggc	900
ccattgccca	cccaccatc	tcactatgca	attcgagttc	caagcaacat	ttgctcctgc	960
cttggggcca	gctctgccc	agccctgaga	ggggtggtga	ggcagcccc	tggaccccag	1020
aaccccagac	aagggggcag	gcgggggacc	agggcctctc	ctgtgggac	tttgttttgt	1080
gtttaaccat	aatggttgtg	tactgaacca	cttcataatt	gttatatata	atatatata	1140
atataatctc	cttaagactc	agcctcctgg	tttaccctcc	cggcctgggc	atctgacctc	1200
cagcctctgt	agggccatgg	ctgtatgtac	tgctgctgtg	tttttttgtt	tttttagaac	1260
tgggtttggg	ggctgatttt	tatttctttg	ggggcttttt	ttcttgga	ataactaaaa	1320
tctcgtcaat	gtaatttctg	tggtttctat	tcagcttggg	tttcatgttt	taaaataaat	1380
tttaaaaagc	aaaaaaaaaa	aaaaaaaaaa				1409

<210> 83

<211> 714

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (704)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (709)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (714)

<223> n equals a,t,g, or c

<400> 83

48

attcggcacg	agaccaaaga	gaggcttggg	acagacccta	ccctctcagt	cttcacagga	60
caacaaccct	gcccgacact	tgatcttggg	cttctagctt	ccagagctgt	gagaaaataa	120
atgtctattg	tttaagccat	ccagtttgtg	gtactttatt	acaacaaccc	tagcaaaacta	180
gtacaaaggg	catgcaattc	attctcactg	gaataacact	gtctgggtat	ttgtttacct	240
tttctgcctg	tgagtgcta	tcagccagca	tcactgtctg	gggtcttatg	gagtgccctga	300
ttcacaggca	tgatccccac	acaacagagc	atctgaccag	gacacttact	tcacagcaaa	360
gctccagagg	tcacctgtcc	ctatcacaca	gcaccacca	aagtaatcaa	cctgaaagaa	420
cattagcact	gcttactgga	ggaacagctg	acttatcggt	ttggaggcaa	cactctccaa	480
agatgggtgc	catcttccaa	gatgctgtat	ttgcacttga	ctcacaggcc	tatttatggg	540
gcattgtttc	taatagggag	aatatatggg	tcctggaaca	atggccccc	cccaaggggt	600
tccattcttg	ccaggaaaca	ccgcaagagt	cccattgagg	ctgggcgtgg	tggtacacac	660
ctgtaatccc	agctactcga	gagtaccttc	taggagcggg	cgnggggcnc	atcn	714

<210> 84

<211> 1097

<212> DNA

<213> Homo sapiens

<400> 84

ccacgcgtcc	gggcgtgct	ttttgcacgt	tctttgcgct	tgtgccgctg	gggagccaaa	60
cgattgggag	ttgcctccac	agaggcccag	agaggcgtca	gtttcaaact	ggaagaaaaa	120
accgcccaca	gcagcctggc	actcttcaga	gatgatacgg	gtgtcaaata	tggtttggtg	180
ggattggagc	ccaccaaggt	ggccttgaat	gtggagcgct	tccgggagtg	ggcagtggtg	240
ctggcagaca	cagcggtcac	cagtggcaga	cactactggg	aagtgcagct	gaagcgctcc	300
cagcagttcc	ggataggagt	ggcagatgtg	gacatgtccc	gggatagctg	cattggtggt	360
gatgatcggt	cctgggtggt	cactatgccc	agcgcaagtg	gtacaccatg	ttggccaacg	420
agaaagcccc	agttgagggg	attgggcagc	caagaagtgg	ggctgttgct	ggagtatgag	480
gccagaagc	tgagcctggg	ggatgtgagc	caggctctctg	tggttcacac	gctacagaca	540
gatttccggg	gtccagtggg	gcctgccttt	gctctctggg	atggggagct	gctgacccat	600
tcagggcttg	aggtgcccga	gggcctctag	tatgtccatt	actggagctc	ctaatacgc	660
ctttggccag	cctccttttg	aaagtgtccg	aagccttttt	actttgcctc	aagcaacctc	720
tagctcccac	aattcagtg	tggtcctct	gtgcaatata	atgatcatct	tcctcatccc	780
ctaccttggtg	aaagctaggc	atacagccaa	accctccttt	tcccaccca	ccaacactac	840
tgccaatttc	ctaggctacc	atgggtgtat	cttccttgac	ctgcttcctt	cagtccctct	900
gcctcccttt	gccaggcct	ttctcagact	gtattccatc	ctgggggtctt	atcattcagc	960
tttgtttgaa	tttattaatc	accatgatac	ctctccctcc	ctttgtccac	atgtaacttg	1020
ttcttggggc	tctaccagat	ggctgaagag	taaatccttt	ctacctctga	aaaaaaaaaa	1080
aaaaaaaaaa	aaaaaaa					1097

<210> 85

<211> 1931

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1904)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1914)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1921)

<223> n equals a,t,g, or c

<400> 85

ggcacgagcg	gcacgagcgg	atcctcacac	gactgtgata	cgattctttc	cagcggcttc	60
tgcaaccaag	cgggtcttac	ccccggctct	ccgcgtctcc	agtcctcgca	cctggaaccc	120
caacgtcccc	gagagtcccc	gaatccccgc	tcccaggcta	cctaagagga	tgagcgggtg	180
tccgacggcc	ggggcagccc	tgatgctctg	cgccgccacc	gccgtgctac	tgagcgtca	240
gggcggaccc	gtgcagtcca	agtcgcccg	ctttgcgtcc	tgggacgaga	tgaatgtcct	300
ggcgacgga	ctcctgcagc	tcggccagg	gctgcgcgaa	cacgcggagc	gcacccgcag	360
tcagctgagc	gcgctggagc	ggcgctgag	cgcgctgggg	tccgcctgtc	agggaaaccga	420
ggggtccacc	gacctcccg	tagccctga	gagccgggtg	gacctgagg	tccttcacag	480
cctgcagaca	caactcaagg	ctcagaacag	caggatccag	caactcttcc	acaaggtggc	540
ccagcagcag	cggcacctgg	agaagcagca	cctgcgaatt	cagcatctgc	aaagccagtt	600
tggcctcctg	gaccacaagc	acctagacca	tgaggtggcc	aagcctgccc	gaagaaagag	660
gctgcccag	atggcccagc	cagttgacct	ggctcacaat	gtcagccgcc	tgacccggct	720
gcccagggat	tgccaggagc	tgttccaggt	tggggagagg	cagagtggac	tatttgaaat	780
ccagcctcag	gggtctccgc	catttttgg	gaactgcaag	atgacctcag	atggaggctg	840
gacagtaatt	cagaggcgcc	acgatggctc	agtggacttc	aaccggccct	gggaagccta	900
caaggcgggg	tttggggatc	cccacggcga	gttctggctg	ggtctggaga	aggtgcatag	960
catcacgggg	gaccgcaaca	gccgcctggc	cgtgcagctg	cgggactggg	atggcaacgc	1020
cgagttgctg	cagttctccg	tgacactggg	tggcgaggac	acggcctata	gcctgcagct	1080
cactgcaccc	gtggccggcc	agctgggcgc	caccaccgtc	ccaccagcg	gcctctccgt	1140
acctttctcc	acttgggacc	aggatcacga	cctccgcagg	gacaagaact	gcgcaagag	1200
cctctctgga	ggctggtggt	ttggcacctg	cagccattcc	aacctcaacg	gccagtactt	1260
ccgctccatc	ccacagcagc	ggcagaagct	taagaaggga	atcttctgga	agacctggcg	1320
gggcgcgtac	taccgcgtgc	aggccaccac	catgttgatc	cagcccatgg	cagcagaggc	1380
agcctcctag	cgtcctggct	gggcctggtc	ccaggcccac	gaaagacggt	gactcttggc	1440
tctgcccag	gatgtggccg	ttccctgcct	gggcaggggc	tccaaggagg	ggccatctgg	1500
aaacttggtg	acagagaaga	agaccacgac	tggagaagcc	ccctttctga	gtgcaggggg	1560
gctgcatg	ttgcctcctg	agatcgaggc	tgcaggatat	gctcagactc	tagaggcgtg	1620
gaccaagggg	catggagctt	cactccttgc	tggccaggga	gttggggact	cagagggacc	1680
acttggggcc	agccagactg	gcctcaatgg	cggactcagt	cacattgact	gacggggacc	1740
agggcttg	tgggtcgaga	gcgcccctcat	ggtgctggtg	ctgttggtg	taggtccctt	1800
ggggacacaa	gcaggcgcca	atggtatctg	ggcggagctc	acagagttct	tggaataaaa	1860
gcaacctcag	aacaaaaaaaa	aaaaaaaaaaa	aaaagggcgg	ccgncctaaa	aggntccaag	1920
nttacgttac	g					1931

<210> 86

<211> 1092

<212> DNA

<213> Homo sapiens

<400> 86

aggccatgac	ctccctcagg	atgcctggct	gcgctgggtg	ctggctgggg	cgtgtgtg	60
cgggtggctg	gcagtgaact	acctcccg	cttcctgatg	gagaagacac	tcttctctta	120
ccactacctg	cccgcaactca	ccttccaaat	ccttctgctc	cctgtggtcc	tgacgacat	180
cagcgaccac	ctgtgcaggt	cccagctcca	gaggagcatc	ttcagcgccc	tgggtgggtg	240
ctggtactcc	tccgcgtgcc	acgtgtccaa	cacgctgcgc	ccactcacct	acggggacaa	300
gtcactctcg	ccacatgaac	tcaaggccct	tcgctggaaa	gacagctggg	acatcttgat	360
ccgaaaaacac	tagaacaaga	gtgtggcaaa	gaacaccg	gctggggtcg	ggacgaggtt	420
gaagggtctt	ggtcaatgta	cgtaatgagc	aggggtggcc	ccacgctggg	aggacacggg	480
ctggctgag	cagggcctct	agtggaaacac	atgggggtct	cattgaaaag	ctctctgatg	540
agcacctcct	tttgtcaaa	gttaattttt	tctcgacaat	aaagatatc	cgtgtcttca	600
cccctgaact	aagacacagg	gagtatttca	gaaggccaag	cgtaggagtc	atcgacaacg	660
aaaaagccga	gaaccacagg	ccagcagttg	gagcctcag	cagaaccagg	gcctggctcc	720
tgctaattgc	tgacgggtg	agtttgatct	ggcagaccg	atcctccttc	atgaacacct	780
agcaacctga	gcaagtcccg	gccctgccct	cagcgagccc	ggcaggcgctc	ccgggacagc	840
tcagtgttg	agggccacct	gaaccacgag	ccagggtctg	ggcttgcatg	tcattgtcta	900
tgacagcgtc	aagactggcc	cttggcaccg	tgctgtgtg	aaacctccc	ctctgagact	960
ccactgagac	gtggctgagt	gaaatcttcc	tcgtcagtg	tcaagggtg	tcattccatac	1020
agtcacatgc	ctttgtcttt	tttaaatgta	attaaaaaag	gaaccaactg	gaaaaaaaa	1080

aaaaaaaaaa aa

1092

<210> 87

<211> 578

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (576)

<223> n equals a,t,g, or c

<400> 87

gggacatctg	ccggctggag	cgggcagtg	gccgcgatga	gccctctgcc	ctggcccg	60
cccttacctg	gaggcaggca	agggcacagg	ctggagccat	gctgctcttc	gggctgtgct	120
gggggcccta	cgtggccaca	ctgctcctct	cagtcctggc	ctatgacag	cgcccgccac	180
tgsggccttg	gacactgttg	tccctcctct	ccctaggaag	tgccagtgc	gcggcagtg	240
ccgtagccat	ggggctgggc	gacagcgc	acacagcccc	ctggagggca	gccgccc	300
ggtgcctgca	ggggctgtgg	ggaagagcct	cccgggacag	tcccggcccc	agcattgcct	360
accaccaag	cagccaaagc	agtgtcgacc	tggacttgaa	ctaaaggag	ggcctctgct	420
gactcctacc	agagcatccg	tccagctcag	ccatccagcc	tgtctctact	gggccccact	480
tctctggatc	agagaccctg	cctctgtttg	accccgccact	gactgaataa	agctcctctg	540
gccgtttaa	aaaaaaaaaa	aaaaaaaaaa	gggggncc			578

<210> 88

<211> 699

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (661)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (694)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (696)

<223> n equals a,t,g, or c

<400> 88

tgcacccacg	cgtccggaag	cccccaacag	ccacgctcac	cactgctcgg	acgaggccga	60
ccacagacat	gagtgcaggt	aagtggctcc	tgctgggtgat	cttcagggat	ttgggatgcg	120
gagtttccag	gacgtctccg	cacttgagga	gtggagagga	gggaaggatc	tggaagcctac	180
tcacagcctg	ctcctgctgt	tgctctctcg	tgatcttcta	gtggttcttg	gcgaaatcag	240
gaaaaggcag	atggagggtt	gtgtatggaa	aggggtggga	tggaatccgg	agaaatggtt	300
tgcggtcttg	gctctgcctg	taacaaccgg	agtgaacctg	ggcaagtccc	tgtccctctc	360
tgggsctcag	tttctccacc	tgtatttgga	raggggttga	atgggcaactg	aagtcctgtc	420
cagctctgac	cttctgtgaa	gtgcactgtt	gagcagctct	ggaagcttct	gttccagcca	480
tagccacaca	gaggagcagc	aggcaggcat	caggccccaa	ctgctgctct	ctgatgggct	540
tggaccccat	gaaagtgggg	cctgctggat	gcatttccctg	ggattctgtg	gaagctgatc	600
aggttgctgg	ggcaagtggg	ggcaggatag	aagtgaaggg	ctgtgggatg	gagaacctca	660
naagactcca	tctgggggtcc	gggaaaggac	agananggt			699

<210> 89

<211> 1126

<212> DNA

<213> Homo sapiens

<400> 89

ggcasagcca	accctgagga	ctcagtgtgc	atcctggaag	gcttctctgt	gactgcactt	60
agcattcttc	agcacctggg	gtgccacagc	ggagcagttc	gtctccctat	tactgtcagg	120
agtgggggca	gattctgctg	ctggggaagg	aaacaggagc	ctggttcaca	gyttagtgat	180
ggagatatga	cctcagccct	aaggggggtt	gctgatgacc	aaggacagca	cccactgttg	240
aagatgcttc	ttcacctgtt	ggctttctct	tctgcagcaa	caggtcacct	tcaagccagt	300
gtcctgaccc	agtgccttaa	ggttttgggt	aaattagccg	aaaacacttc	ctgtgatttc	360
ttgcccagg	tccagtgtgt	gttccaagt	ctgccaaagt	gcctcagccc	agagacaccc	420
ctgcctagcg	tgctgctggc	tgttgagctc	ctctccctgc	tgccggacca	cgaccagctg	480
gcacctcagc	tctgttccca	ctcagaaggc	tgccctcctgc	tgctgctgta	catgtacatc	540
acatcacggc	ctgacagagt	ggccttgagg	acacaatggc	tccagctgga	acaagagggtg	600
gtgtggctcc	tggttaagct	tggtgtgcaa	gagccccttg	cccccagtca	ctggctccaa	660
ctgccagtgt	aatgtggagg	tggtcagagc	gctcacgggtg	atgttgacaca	gacagtggct	720
gacagtgcgg	agggcagggg	gacccccaa	gaccgaccag	cagaggcgga	cagtgcgctg	780
tctgcgggac	acgggtgctg	tgctgcacgg	cctatcgagc	aaggacaagc	tcttcatgat	840
gcactgcgtg	gaggtcctgc	atcagtttga	ccaggtgatg	ccgggggtca	gcatgctcat	900
ccgagggtct	cctgatgtga	cggactgtga	agaggcagcc	ctggatgacc	tctgtgccgc	960
ggaaaccgat	gtggaagacc	ccgaggtgga	gtgtggctga	ggccctgagt	gtccagccac	1020
atggtggcac	cagcaccact	cctttcctta	ccacatcaac	tgattaaagc	agtgaccagc	1080
aggaactgcc	cagagaactg	gaaaaaaaa	aaaaaaaaa	ctcgag		1126

<210> 90

<211> 1037

<212> DNA

<213> Homo sapiens

<400> 90

agggttgatg	ggttatggtc	aggagtccca	gctgggcccc	ccacctcctc	aggaaggcgg	60
gtgaggttgg	tgtgagactg	acgggtgcctc	ctcatgtccc	cttggagcgc	cccacccccac	120
atctcccggc	ctcgggtcct	tgcttgcccc	agcatgagag	gtgcttcata	ggaacggagg	180
gaggacatgt	ygggacagct	cgatgctcgg	cctgctgctg	ctctgcaccc	ccagggcctg	240
gctcacccctc	tctggacctg	tctgcttcca	aggaagggga	ccctctgagg	tcccacagag	300
gccacccccag	ytgtgggtcg	tgagcatctc	tgtcttgtag	ggacagcatc	gtggccgagc	360
tggaaccgaga	gatgagcagg	agcgtggagc	tgaccaacac	camcttcctg	ctcatggccg	420
cctccatcta	tctccacgac	cagaaccccg	atgccgccct	gcgtgcgctg	caccagggggg	480
acagcctgga	gtggtgagtg	gcctccctgc	tctgggccag	cccaggaggag	caagtgcctc	540
ctgccacatc	tccaggctgc	gcacggcctc	gctggctgtc	gtcatgggag	cagagaaagg	600
tggtgctgaa	atgaggccct	ggcctgctgt	ccaggctcca	gctccctcgc	ccagtgtggg	660
aggcactccc	atctgcgcac	cagggtgcgg	atccaaggac	acggtgccca	rgctgcaacc	720
ctctgttccc	aagggcagag	cagaaaagcg	ctttgtctct	gctcggtttc	tgtgtcccca	780
ccccccacga	agccttctgt	gtctcggccc	tgggcccagt	ctctcaggcc	tccccgggcc	840
ccccataccg	gccctcctcc	agggccctct	ggggttgggg	tgctgaagcc	ctgcaagggtt	900
ggtgcccccc	tccaccctag	gatgtgactc	cgggccatgt	ccagggcact	ggtcacagaa	960
agtgtgtcag	ttcttccccg	tgagctgtcc	ctgcagtgcc	tgccctccac	tgtgagttgc	1020
aagctgggca	tttcatg					1037

<210> 91

<211> 1316

<212> DNA

<213> Homo sapiens

<400> 91

ggcacgaggc	ctggcgcgct	gcggcgctgc	tcacccgctc	ccgaggaagg	gcagtggggc	60
ccgccgcgcg	ctcccaatgg	cgaggctgcg	ggattgcctg	ccccgcctga	tgctcacgct	120
ccggctccctg	ctcttctggg	ccctggtcta	ctgctactgc	gggctctgcg	cctccatcca	180

52

cctgctcaaa	cttttgtgga	gcctcggcaa	ggggccggcg	cagaccttcc	ggcgcccgcc	240
ccgggagcac	cctcccgcgt	gcctgagcga	cccctccttg	ggcaccact	gctacgtgcg	300
gatcaaggat	tcagggttaa	gatttcacta	tgttgctgct	ggagaaagag	gcaaaccact	360
tatgctgctg	cttcatggat	ttccagaatt	ctggtattct	tggcgttacc	aactgagaga	420
attttaaagt	gaatatcgag	ttgtagcact	ggatttgaga	ggttatggag	aaacagatgc	480
tcccattcat	cgacagaatt	ataaattgga	ttgtctaatt	acagatataa	aggatatttt	540
agattcttta	gggtatagca	aatgtgttct	tattggccat	gactgggggg	gcatgattgc	600
ttggctaatt	gccatctgtt	atcctgaaat	ggtgatgaag	cttattgtta	ttaacttccc	660
tcattccaaat	gtattttacag	aatatatttt	acgacaccct	gctcagctgt	tgaaatccag	720
ttattattac	ttcttccaaa	taccatgggt	cccagaattt	atgttctcaa	taaatgattt	780
caaggttttg	aaacatctgt	ttaccagtca	cagcactggc	attggaagaa	aaggatgcca	840
attaacaaca	gaggatcttg	aagcttatat	ttatgtcttt	tctcagcctg	gagcattaag	900
tggcccaatt	aaccattacc	gaaatatctt	cagctgcctg	cctctcaaac	atcacatggt	960
gaccactcca	acactactac	tgtggggaga	gaatgacgca	ttcatggagg	ttgagatggc	1020
tgaagtcaca	aagatttatg	ttaaaaacta	tttcaggcta	actattttgt	cagaagccag	1080
tcattggctt	cagcaagacc	aacctgacat	agtgaacaaa	ttgatatgga	cattttctaaa	1140
agaagaaaca	agaaaaaaag	attgactttt	ctttatcttc	tatgaagggt	ctgtaatgaa	1200
atctctaaat	aatttttaaa	aattgttcat	caacttcttt	atgttttatt	agaaaaaac	1260
tgttttaatg	tgctttatca	taaataaata	tcctgacaaa	tggtattgaa	aaaaaa	1316

<210> 92

<211> 1021

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (971)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1004)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1008)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1010)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1018)

<223> n equals a,t,g, or c

<400> 92

ggcgcgccctt	tttttttttt	tttttttttt	tttttttttt	ttttggcctt	agtcattcatt	60
tcttgaataa	tacaaatagg	taagacaatt	ttacaaaaat	tgtgctatag	aataggatat	120
ttgtgacttt	ttagatgaaa	tattagagct	accccaccca	gccacagata	gcaactgtaac	180
actttcttaa	tagagtatat	gttcaaatta	taaagtccac	acactggcta	aaaagttcaa	240
gttcagagtt	tcaatcaatt	ttcattgtaa	ggatgaaact	gagttttact	caacttgtgt	300
ctttttaaga	gaatgggcca	cctcccacac	atcctttctc	ttggactttt	tttaacactt	360
ctaagtgtct	gtatcacgaa	atcagatggc	caaaacaaaa	tctacagggt	ctttaaaaaa	420
gcaagtcgcc	aagtgattgt	taccatatac	aaaatgagaa	ttgctgctat	aatctgttct	480

53

tactggamtg gccakgccaa tcttgggact aggattaaat tgcaattaaa ttckgcagtg	540
tacaaaattt ttgtcagtc gyctagaaaa agaaagagaa ctctttcatg gtagagcagt	600
tactgtgctc acgttgcttt ttctaaaaac caacctactt tcaaacaaag aatgaggaaa	660
tttgagtaa attttaaata tgagtcacgg aaatattaag ataatagcat gtgtgggcaa	720
taataagtat gccaaagaaat aaagagtaat atacaaaaca atcaaacatt attacatttg	780
gctacgaggt tcctaataaa cagggcaaaa taaatagtga aatataataa aatcgttatc	840
atctgataaa aggctgcatg gtacttttcc caaacgtaat ggatgacttc aacacatttt	900
cttattaaat atttcaaatt gtttcttcat gtgaaaactg tcttattaat tgtaaaaagg	960
atgtaacttg nataggcatg ctcaacaggg gtaagagtaa ttcngtangn gccccctnga	1020
t	1021

<210> 93

<211> 1260

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (32)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (314)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (356)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (590)

<223> n equals a,t,g, or c

<400> 93

tttttttttt tttttttttt tttttttttt tntttttttt tttttttttt tttttttttt	60
tttttttttt ttttttttta aatttcacct gtttccttta ttatgtggct acttgaaaaat	120
tttaaaatta catatgtagc tcacactata aaacacagat tagaaatatt gtatagcact	180
gacctagaaa cctccattta ggtaaaacat cttaaccctt ttggaagcaa aatatgttaa	240
ataacagcat aaactccac caagaaaatc ctcaccttcc tcctttcaac acatttatta	300
tatacagctg tcantgcatt gtcaatctgc caaatggctc tatgttccaa cagggntgga	360
gtagtccctt gctcacacca gccttcacaa tacttcccat gtcttccctg ttaacctctc	420
tccaccagc acccaggctc ccaactctcc tggctgcctc cagccctcag ctggcaccac	480
tgacatgctg tttccagtac ccttttcttc tttctgcatc ctccctgggg gacatacatc	540
cctcatctcg tgacttcagc tgtcacataa attcaaagt ttccagaactn tatttttttac	600
ctctacatc tgtcagttta aatgtcagga tatttttact tcagtaaagc cctaaaaaga	660
caaatctatg tactttttaa gaataaaaga aatgactggc tgcagctcaa acctacaact	720
gcttgcgaaa ctctacaatg tctggcagat gctagaaaaga aggggatcaa gacagagcac	780
acttggcgtg gtatgctatc tatagaaaat gttaaaataa aattaagtaa tctaggtttc	840
ctctctttat tttctacatc tactctctga agagggcaat aaataaggaa atgtcccaaa	900
gagggacaaa ttaagtccca aaataacaca aaattgggca aatcccagtc atgaagaaag	960
aacagagggt cttaaattgg gacacacaga ggcaggctcg caggtctagg aatctctgaa	1020
catatgtgca aaattctggg tatgtgtgca tatgttatat aacaaagcga aggggtccata	1080
tagctttcat cgcatttcaa aggggtctagc actgaaataa ggactactgc tatgtgactt	1140
aaaaaatgaa actcaggctg ggcgcagtc tcacgcctgt aatcccagca ctttgggagg	1200
ccgaggcaag cagatcacct gagacgagga gtttgagacc cgcttgccg gacgcgtggg	1260

<210> 94
 <211> 990
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (4)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (916)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (958)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (971)
 <223> n equals a,t,g, or c

<400> 94	
gcangagagc taccaagtgg ccgagctggc ttatraccct agggcagtct gcccccaaaa	60
cccatctact tggttgtccc cagaatgggt tggcttggga gaacttgcct tgctcactcc	120
catttagact ttattagtgg agccctctc ttgacttttg cctatttcct tgtctttcag	180
gtgtgccttg tgattaataa atggctctac aacctggacc agcatgtggt taaagagttg	240
attagtaagt gctggagggtg ggaagggaca ggaacactcc agaagaaaag tcagaaccct	300
ccctcaccct ttgtatttca tttcccctta cctcactctg gcacttctcc tagaccaaaa	360
atctctttcc tgctgaagta gaatggtccc taataataac aaccttaata ataaactcag	420
ctgacattaa ctgagggagc ccagtgtgcc aacatgaagc actgtgcctg cactagcaat	480
tgaacgtgca ccttttagcta aggacgtgct ggtttcaatt ctattcttgc tcccaagcct	540
acagcagctg agatatgaat ggaaacttct ccaggggaga aaatctgccc aattctgcct	600
ttgtcctccc ctaaatttgt atgagttaaa tgatgggcag aaaattggtc tgttttcagc	660
ccagacaaac actgcctcct ttcagtagtc gctacctcaa gcatccaaag ttttcatatc	720
tgccagaact caaagcaaaa aatgcaagat tgaatctcag cagctcaggc cccagcagg	780
acttcaaact tccaccacca aaaaaaaaaa aaaaaaaaaa gctgaattga aaggtatatg	840
ccttcattca ctgaatattc actcgtcctg ccaagtgcc gatgccarag tttctaaaat	900
tcccccaaag gggggnccgg gtacccaatt cccccctatt agtgaagtcc tatttacnaa	960
ttcccttggtg nccgtccgtt tttaacaacc	990

<210> 95
 <211> 1710
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1702)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1704)
 <223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1709)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1710)

<223> n equals a,t,g, or c

<400> 95

ccaggaattc	cggggtcgac	ccacgcgtcc	ggaaacattt	cccatgtcct	aagttcttag	60
aagcaattac	tttagcgttg	gggagcattg	ttctaccaga	cccatttatg	caggggagat	120
gaagcttaaa	agacgggatg	tggggtggga	gtgtgtttct	taagccgaag	ctgttgcagg	180
ctgggggatt	tttgcathtt	ctttttgttt	tgtttttgac	tgcagattct	gtacatctgt	240
ctgtgggagg	agagttgcta	ctcaggacag	gatttaagag	acacattcca	gtgaccttta	300
agaatctgca	tggcgggagg	tccttctcca	ggagtgtggg	ttgggtccact	ctgggaccca	360
ccacactaag	aagggggaga	tgataaaata	acattaaagg	aagaatggcc	tccagcctgc	420
aggttttgtt	ggaaagaaat	aaaaaggagg	tcattaagac	cataaattca	gattgaggcc	480
tcttgaaaag	gttgatgtgg	gctagcaacc	tgctgtcga	aacagtcctt	ggttcatcag	540
cagtgtggga	aggcagccaa	ggctcctgca	gattcctggc	atcgaccttg	gaaagcctct	600
gcgatacttg	tgtgtacgaa	tacagcagag	gacaggagg	tccttgttct	gtggtctctg	660
tttagtgact	gaaacttaaa	cccaaaggca	agcccagatt	gtctggccgt	tcaccccat	720
gctttgatag	gggttaggag	gaaccctttc	cgtatgaaag	acaggcccta	ytagggytta	780
caaccaagcc	aaaggaccat	ctcttctttc	ttccacctcc	cttcamccct	gccccgcagc	840
agagccgaga	tgtgagacat	tcattgtcac	ggagcaagga	gacaaaggca	agttcaagtt	900
gagaagcata	tggcagcaaa	cagaaatgaa	aaccatatgt	cccagcaagg	ggaaaagcag	960
tcatttccag	attataaaaa	tcaatgaagt	actctccacc	taggtcagct	gaaattcgag	1020
ccctcacagt	caggcctgtc	agagaagtta	agcagaaaca	tctcgggggg	acttctaaaa	1080
tttagtgaag	acaaggcctt	gcaactccaa	agaaactttt	tttccccccc	ttgaaacagg	1140
gtcttgctct	gttgcccagc	ctggagtgc	gtggtgcagt	cacggctcac	tgcagcctca	1200
agtccttggg	ctcaagcact	atccccacct	caacctcctt	agtagctggg	actacaggtg	1260
cacaccacca	tgcccagcta	accacagaaa	ctttcatctc	ttcatttttt	ctttgggcac	1320
cattaatacc	taagacaggt	agaaagggtc	ccagaaagac	accattggta	atggccgatt	1380
gccggctgca	gtcatcgccc	ccagatcagg	ctggtacagg	atgccttaag	gtgatgagag	1440
gtgagggtgc	atgaagaata	atgagcacag	ggaagagaga	agcaggacaa	agtagcagat	1500
aaaatgccgg	caaagcacag	atgaatgtct	tcaagaagct	cttgattttc	tctgcacagt	1560
gtaaatatcc	ttgctatttc	aggatggcgg	ctggcctgct	cagtaacata	catgttccaa	1620
ataaagattt	tgcattgaaag	taaaaaaaaa	aaaaaagggc	ggccgctcta	gaggatccaa	1680
gcttacgtac	gcgtgcatgc	gncngtcann				1710

<210> 96

<211> 781

<212> DNA

<213> Homo sapiens

<400> 96

cggcacgagg	cagccagtag	gggagagagc	agttaaggca	cacagagcac	cagctccctc	60
ctgcctgaag	atgttccacc	aaatttgggc	agctctgctc	tacttctatg	gtattatcct	120
taactccatc	taccagtgcc	ctgagcacag	tcaactgaca	actctggggc	tgatgggaa	180
ggagtcccca	tgggtccact	tgggcccagt	gtactttatc	gcaggggcag	ctcccaccaa	240
ggaggagtgt	gcaacttttg	accctgtgga	caacattgtc	ttcaatatgg	ctgctggctc	300
tgccccgatg	cagctccacc	ttcgtgctac	catccgcgat	aaagatgggc	tctgtgtgcc	360
ccggaaatgg	atctaccacc	tgactgaagg	gagcacagat	ctcagaactg	aaggccgccc	420
tgacatgaag	actgagctct	tttccagctc	atgcccagg	ggaatcatgc	tgaatgagac	480
aggccagggt	taccagcgt	ttctcctcta	caatcgctca	ccacatcctc	ccgaaaagt	540
tgtggaggaa	ttcaagtccc	tgacttctct	cctggactcc	aaagccttct	tattgactcc	600
taggaatcaa	gaggcctgtg	agctgtccaa	taactgacct	gtaacttcat	ctaagtcccc	660
agatgggtac	aatgggagct	gagttgttgg	agggagaagc	tggagacttc	cagctccagc	720

56

tcccactcaa gataataaag ataatttttc aatcctcaaa aaaaaaaaaa aaaaactcga 780
g 781

<210> 97

<211> 1113

<212> DNA

<213> Homo sapiens

<400> 97

gaagatttgg gagcatctga agagccagag gagttagagg ctctgaagca cagtgacttg 60
atgtctaagc tgtttctttg ttcaccccc tggttggtt aaatctaarc tgtctctttg 120
cttgtatgat catagtctcc tgcacatcggg ttataagatt ctccaaattc agaatgcgca 180
ttgcaaattc ctcatggtag gctggtgaag gacttgacaa tctgatgata aattactttg 240
tgaacatgaa tgaaaaatat tgcacatcat attggtgagg agaaagtgga gtaaagaaaa 300
atctccatct atacacaata cagcttttta aatgcacggg actttcaaatt atttgcatct 360
ctacattata cgttttgttt caacttacgc atttattgtt ttctttcctt tttcttcctc 420
acatgttaat ggcccatgtg agaaaaacat tcccctgggt aaatagatag aagagatttg 480
tgcaaattgca agagaaattt cagtgtatct gctatgattt gaatgtgtcc cccaacgttc 540
atgtgttgca aatttgattc ccaatgcagt ggtgttgagg agtgaggcct aatggaagg 600
gtttgggtca tgggggcacc gccttcataa atggattaat gccattattg tgggaatggg 660
ttccttataa aaagatgagt tcagtcacct cttgctctcc tctcacctc tctttgccct 720
tttaccatag gctgacacag caagaaggct cttgccagat gctggtacct tgatattgga 780
tttcccaggc tccagaacta aaaagaatga atttctttt tttttaaat acccagctctg 840
tggtaaattt atagtagcac aaaacagatt aagacaatat gtctttcaga tgtcttagct 900
tatttcattg gataactgta agaagagttt actgctttt ttttttgaa attaagaatt 960
tagctgaatg ctgttgctca cacctgtaat tccgcactt tgggagcgagg aggcgggctg 1020
atcacctgag gtcaggagtt tgagaccagc ctggccaacg tggtgaaact ctgcctctac 1080
taaaaatata aaaaaaaaaa aaaaaaactc gta 1113

<210> 98

<211> 1723

<212> DNA

<213> Homo sapiens

<400> 98

gaattcggca cgagcgacat gggctccgct ccctgggccc cggctcctgct gctggcgctc 60
gggctgcgag gcctccaggc ggggggtgag tggcgcgccc ccccggccca tccccggctc 120
ccggccccgc ctctgagggt cgctccccc cacagcccgc aggcgccgga ccccggttc 180
caggagcgct tcttccagca gcgtctggac cacttcaact tcgagcgctt cggcaacaag 240
accttcccyt agcgcttctt ggtgtcggac aggttctggg tccggggcga ggggccatc 300
ttcttctaca ctggaacga gggcgacgtg tgggccttcg ccaacaactc gggcttcgtc 360
gcggagctgg cggccgagcg gggggtctta ctggtcttcg cggagcaccg ctactacggg 420
aagtcgctgc cgttcgggtgc gcagtcacg cagcgcgggc acacggagct gctgacgggt 480
gagcaggccc tggccgactt cgcagagctg ctccgcgcgc tacgacgcga cctcggggcc 540
caggatgccc ccggcatcgc cttcgggtga agttatggg ggatgctcag tgcctacctg 600
aggatgaagt atccccacct ggtggcgggg gcgctggcgg ccagcgcgcc cgttctagct 660
gtggcaggcc tcggcgactc caaccagttc ttccgggacg tcacggcgga ctttgagggc 720
cagagtccca aatgcacca ggggtgtcgg gaagcgttcc gacagatcaa ggacttgctc 780
ctacaggag cctacgacac ggtccgctgg gagttcggca cctgccagcc gctgtcagac 840
gagaaggacc tgaccagct cttcatgttc gcccggaatg ccttcaccgt gctggccatg 900
atggactacc cctacccac tgacttcttg ggtccctcc ctgccaaccc cgtcaagggtg 960
ggctgtgatc ggctgtgag tgaggcccag aggatcacgg ggctgcgagc actggcaggg 1020
ctggtctaca acgctcggg ctccgagcac tgctacgaca tctaccggct ctaccacagc 1080
tgtgctgacc cactggctg cggcaccggc cccgacgcca gggcctggga ctaccaggcc 1140
tgacccgaga tcaacctgac cttcggcagc aacaatgtga ccgatatgtt ccccgacctg 1200
cccttactg acgagctccg ccagcgggtac tgcctggaca cctggggcgt gtggccccgg 1260
cccgactggc tgctgaccag cttctggggg ggtgatctya gagccgccag caacatcatc 1320
ttctccaacg ggaacctgga cccctgggca gggggcggga ttcggaggaa cctgagtgcc 1380
tcagtcacgt ccgtcaccat ccagggggga gcgcaccacc tcgacctcag agcctccac 1440

57

ccagaagatc	ctgcttccgt	ggttgaggcg	cggaagctgg	aggccaccat	catcggcgag	1500
tgggtaaagg	cagccaggcg	tgagcagcag	ccagctctgc	gtggggggcc	cagactcagc	1560
ctctgagcac	aggactggag	gggtctcaag	gctcctcatg	gagtgggggc	ttcactcaag	1620
cagctggcgg	cagagggaag	gggctgaata	aacgcctgga	ggcctggcma	aaaaaaaaaa	1680
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaggcgccg	cgc		1723

<210> 99

<211> 2087

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> n equals a,t,g, or c

<400> 99

tcgaccacag	cgtccgtggg	gccgagcgcc	gctgggtagg	cggaagtagc	cgcagnatgg	60
cggcggttat	gccgcttgct	ctgctcgtcc	tggtgctcct	ggggcccgcc	ggctggtgcc	120
ttgcagaacc	cccacgcgac	agcctgcggg	aggaaactgt	catcaccccg	ctgccttccg	180
gggacgtagc	cgccacattc	cagttccgca	cgcgctggga	ttcggagctt	cagcgggaag	240
gagtgtccca	ttacaggctc	tttcccaaag	ccctggggca	gctgatctcc	aagtattctc	300
tacgggagct	gcacctgtca	ttcacacaag	gcttttggag	gacccgatac	tgggggccac	360
ccttcctgca	ggcccatca	gacactgacc	actactttct	gcgctatgct	gtgctgccgc	420
gggaggtggg	ctgcaccgaa	aacctcacc	cctggaagaa	gctcttgccc	tgtagttcca	480
aggcaggcct	ctctgtgctg	ctgaaggcag	atcgcttggt	ccacaccagc	taccactccc	540
aggcagtgc	tatccgccct	gtttgcagaa	atgcacgctg	tactagcatc	tcctgggagc	600
tgaggcagac	cctgtcagtt	gtatttgatg	ccttcacac	ggggcaggga	aagaaagact	660
ggtccctctt	ccgatgttc	tcccgaaccc	tcacggagcc	ctgccccctg	gcttcagaga	720
gccgagtcta	tgtggacatc	accacctaca	accaggacaa	cgagacatta	gaggtgcacc	780
cacccccgac	cactacatat	caggacgtca	tcctaggcac	tcggaagacc	tatgccatct	840
atgacttgct	tgacaccgcc	atgatcaaca	actctcgaaa	cctcaacatc	cagctcaagt	900
ggaagagacc	cccagagaat	gaggccccc	cagtgcctt	cctgcatgcc	cagcggtagc	960
tgagtggcta	tgggctgcag	aagggggagc	tgagcacact	gctgtacaac	acccacccat	1020
accgggcctt	cccgggtgctg	ctgctggaca	ccgtaccctg	gtatctgcgg	ctgtatgtgc	1080
acaccctcac	catcacctcc	aagggcaagg	agaacaaacc	aagttacatc	cactaccagc	1140
ctgcccagga	ccggctgcaa	ccccacctcc	tggagatgct	gattcagctg	ccggccaact	1200
cagtcaccaa	ggtttccatc	cagtttgagc	gggcgctgct	gaagtggacc	gagtacacac	1260
cagatccata	ccatggcttc	tatgtcagcc	catctgtcct	cagcgccctt	gtgccagca	1320
tggtagcagc	caagccagtg	gactgggaag	agagtcccct	cttcaacagc	ctgttcccag	1380
tctctgatgg	ctctaactac	tttgtgcggc	tctacacgga	gccgctgctg	gtgaacctgc	1440
cgacaccgga	cttcagcatg	ccctacaacg	tgatctgcct	cacgtgcact	gtggtggccg	1500
tgtgtacagg	ctccttctac	aatctcctca	cccgaacctt	ccacatcgag	gagccccgca	1560
caggtggcct	ggccaagcgg	ctggccaacc	ttatccggcg	cgcccgagg	gtccccccac	1620
tctgattctt	gccctttcca	gcagctgcag	ctgccgtttc	tctctgggga	ggggagccca	1680
agggctgttt	ctgccacttg	ctctcctcag	agttggcttt	tgaaccaaag	tgccctggac	1740
caggtcaggg	cctacagctg	tgttgtccag	tacaggagcc	acgagccaaa	tgtggcattt	1800
gaatttgaat	taacttagaa	attcatttcc	tcacctgtag	tgccacctc	tattattgag	1860
tgctcaataa	gcaaaagtgg	tcgggtggctg	ctgtattgga	cagcacagaa	aaagatttcc	1920
atcaccacag	aaaggtcggc	tggcagcact	ggccaagggt	atgggggtgtg	ctacacagtg	1980
tatgtcactg	tgtagtggat	ggagtttact	gtttgtggaa	taaaaacggc	tgtttccgtg	2040
rwaaaaaaa	aaaaaaaaaa	gggcggccgc	tctagaggat	ccctcga		2087

<210> 100

<211> 751

<212> DNA

<213> Homo sapiens

<220>

<221> SITE
 <222> (663)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (702)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (705)
 <223> n equals a,t,g, or c

<400> 100
 cggcacgagc tttttctggt attccaataa attgtaggtt cagttttttt tatgaagtcc 60
 catatttctt ggaggtcttg ttcattgctt ttaattcttt tttctctaata cttgtctgca 120
 tgctttatctt cggcaagggtg gtcttcaaac tctgatatact tttttctgct ttggctcgatt 180
 cagctattga tacttggtga tgcttcatga agtccccatg ctgtgttttt cagctccatc 240
 aggtctttta tgcttctctc taaactggct attctactta gcaattcctc taaccttttg 300
 tcaagggtct tggtctcttt gtgttggtt aggacatgat cctttagctc agcatagttt 360
 ttcattaccc atcttctgaa gcctacttct gacgttgcca tcatttgagg gagaagaggc 420
 actctgggtct tttgggtttt caaaattttt tcattgtttc tttctcatct ttgtgcattt 480
 gtctagtttc ggtctttgag gccgctgacc ctgggatggg gttttctggt gggctttttg 540
 ttgttcttga tgctgttggt gttgctttct gcttggttgt tttcttttca atggctgggt 600
 ccctcttggt tagggctgct gaagtttgct ggggggtcac ttcaggctct attcatctga 660
 ttactcgca tgcctggaga gtccacttaa gaagcccgga tnaengcata gacagggtgcc 720
 tgctccttct tctgtgatct ctgacctcga g 751

<210> 101
 <211> 1223
 <212> DNA
 <213> Homo sapiens

<400> 101
 gctgctccgt ttttccccca tctttgtggt tttatctacc tttggctttt gatgatgggtg 60
 atgtacagat ggggttttgg tgtggatgtc ctttctgttt gttagttttc cttctaacag 120
 tcaggaccctg cagcttcarg tctgttgagg tttgctggag gtccactcca gacctcttt 180
 gcctgggtat cagcagcaga agctgcagaa cagcggatat tggatgaacag cagatgttgc 240
 tgcctgatcg ttcctctgga agttttgtct cggagtacct agccatgtga ggtgtcagtc 300
 taccctact gggggatgcc tcccagttag gctacttggg agtcaggagc gcacttgagg 360
 aggcactctg tctgttctca gatgtccagc tgtgtgctgg tagaaccagt gctctyttca 420
 aggtkctcag acaggagcgt ttaagtctgc agaggattct gctgcctttt gtttggtgtg 480
 gccctgcccc ccagggttg agtctacaga ggcaggcagg cctccttgaa ttgagggtgg 540
 ctccaccgag ttcgagtttc ctggccgctt tgtttacccc ctcaagcctc ggcaatggtg 600
 ggcgcccctc cccagcctc actgccgctt tgcagtttga tctcagactg ctgtgctagc 660
 aatgakttag gctctgtggg ttagraccc tctgagccag gcatgggata taatctctg 720
 gtgtgctgatt tgctaagacc cctttcttt gactaggaaa ggggtgggaa tgaccaatt 780
 ttccagggtgc cgtctgtcac cccttcttt gactaggaaa ggaattccc tgaccggtg 840
 tgcttcccgg gtgaggcaat gcctcgccct gcttcagctc aagcttgggt cgctgcacc 900
 actgtcttgc acccactttc caacactccc tagtgagatg aaccgggtac ctcagttgga 960
 aatgcagaaa tcacacgtct tctgctcct cagctggga gctgtagact ggagctgttc 1020
 ctattcggcc atcttggtc cacctgtcga gatattttac attaaactttc tatgacatac 1080
 ttatagcaaa acttattttt tcatgcagaa tagtctatat tctatattta ttgtaaagca 1140
 tataccgtac atggtgacta gtcaccatgc tgtacaataa attttctgaa ctttaataaaa 1200
 aaaaaawaaa aaaggcgccg cgc 1223

<210> 102
 <211> 1010

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (607)

<223> n equals a,t,g, or c

<400> 102

ggttacttcc	aagttctgcc	aactgtgaat	aaagttgcta	taaacatcta	tgtacaggtt	60
ttttttgtgt	gtggacctaa	gttttcaact	cctttgggtg	ataccaagga	gcacagtcac	120
tgggacatat	ggtaaggata	tatttagttt	ggcaggaaac	caccatactg	tcttccaaag	180
tagctgtacc	atthttgcata	cccaccagca	ctgaatgaga	gttcctgttg	ctccacattc	240
ttgtcagcat	ttgatgttgt	cagtgttctg	aatttaggta	gtcatgatag	gtgtgtaatg	300
gtatctcact	attatthttaa	tttgcctttc	tctgatgatg	tatgatgttg	cagatcttct	360
catatgctta	tgtgacatct	gtatatctgg	tgaatgtct	gctaagggtc	tascctatth	420
tttaatargg	atggttgtht	tcccattgtt	gagthtttaag	agthccttat	atathtttgg	480
taththaaata	tactacaaat	aaacagtcct	thaacagata	aatgthttgc	aaaththttc	540
tcttagtctg	tggcttctgt	ctthattccc	ttgaagggtg	ctgtcacaaa	gcagthttatc	600
ttthtttctt	ttthtttttt	tttgagacgt	agthctgtct	cagcttggtg	ggcagagcga	660
rctacgtctc	aagaaacaaa	acaaaacaaa	aaaacacctc	agthgcgcgg	caaggthkgt	720
cacgcctgtg	atcccatcac	tttgggaggt	cggaggtggg	aggtggggaga	atcgcttgag	780
gccaggagtc	catcctaggt	ctagcttgac	cctatctcaa	caacaaaaaa	ataacaatta	840
gcccaccgtg	gtagtgcatt	tctgtagtcc	tagctactgg	ggaggctgag	gtgagaggat	900
tgcttgagcc	catgagtht	aggttacagt	gggtataat	tacaccactg	cactccagtc	960
tgagtgcacg	agcaagaccg	tgtctcaaaa	aaaaaaaaaa	aaaactcgag		1010

<210> 103

<211> 1986

<212> DNA

<213> Homo sapiens

<400> 103

ggcacgaggg	aaaactgttt	tatttgcatt	tgaagaagct	attggatata	tgtgctgccc	60
ttttgttctg	gacaaagatg	gagtcagtgc	cgctgtcata	agtgacagat	tggctagctt	120
cctagcaacc	aagaatttgt	ctttgtctca	gcaactaaag	gccatttatg	tggagtattg	180
ctaccatatt	actaaagctt	cctatthttat	ctgccatgat	caagaaacca	thaaagaaatt	240
atthgaaaac	ctcagaaact	acgatggaaa	aaataattat	ccaaaagctt	gtggcaaatt	300
tgaathttct	gccattaggg	acctthacaac	tggctatgat	gatagccaac	ctgataaaaa	360
agctgttctt	cccactagta	aaagcagcca	aatgatcacc	ttcacctttg	ctaattggagg	420
cgthggccacc	atgcgcacca	gtgggacaga	gccccaaaac	aagtactatg	cagagctgtg	480
tgccccacct	gggaacagtg	atcctgagca	gctgaagaag	gaactgaatg	aactggctag	540
tgctattgaa	gaacatttht	tccagccaca	gaagtacaat	ctgcagccaa	aagcagacta	600
aaatagthca	gccttgggtg	tacttgcatt	tacctacaat	taagctgggt	thaaacttgt	660
aagcaatatt	thtaaggggc	aaatgattca	aaacatcaca	ggtatthtat	tgthttacaa	720
agacctacat	tcctcattgt	ttcatgtttg	acctthtaag	tgaaaaaaga	aaatggccaa	780
acccaacaaa	ctaacattcc	tactaaaaag	ttgagcttgg	acatathttg	aaththttgt	840
agtgaagatt	ththaaactga	ctaactthaaa	aaaatagatt	gtaattgatg	tgctthtaatt	900
tgcataaaac	ataaatgtat	gtcctctctg	taattgttht	aatgtgtgct	tgaaatatcc	960
agaaaacct	tggagttagt	aaattctggg	ctgtcatatg	taggatagcc	actthtttag	1020
tatatgtaca	ththaththt	tatcaattcc	thtagaaagta	aaataaatga	atagatcaaa	1080
tgthgtgttc	atgthttggg	aaaatataat	thgcagaaac	ctatgaagta	gagcaaaagat	1140
gctthaaaaa	gataagthtt	ththgaactaa	atthththth	gttctaataa	tgacataggt	1200
atattagtag	atcgtacacg	tgctaggaaa	aaacagcttc	agthgtcttg	ththaatgtgt	1260
tgaaactcat	ctthththaa	cttgaaaaac	caattgttht	cttgaaactt	gaaagtagca	1320
tathththct	thththtggt	gtthgttcat	thgtatttag	acaaththaa	gtaattctct	1380
gtthtgaggc	agcaagacct	atgagcaaga	actaththact	tgacctctgt	ththththct	1440
tgthcttggt	tggcttgaaa	tctaaaaact	gactththth	tgatagattt	cctataagcc	1500
aathththth	aacaaataga	ththththth	aatctgtacc	thctatcttc	tcataattcg	1560

60

tggtcttaca	gccttccaaa	ataactccag	ttgggcaccc	atgagctagg	atcaaaacttt	1620
ctttatatac	tttatatatt	ttacattatt	tctgattttt	aaagcaaagt	attgccatta	1680
tgattacact	caacctaaat	agttatgaac	agtttcagaa	caatgaaaaa	ttacaatact	1740
atgtgatagt	attgtaacta	tttttctatt	ttagtcatat	gtcgccttata	tcctaccaga	1800
actcttaaatt	ctataatatt	cgatatattc	tacaaactgc	tttattgtag	aagccatatt	1860
tatgtttatt	ttataatggt	ttctagtgtc	aaactgtact	gtggagaaaa	gaaatgtag	1920
atctgtgttc	tgtctgcatt	ttttttgagt	acataccctt	cacctcaaaa	aaaaaaaaaa	1980
aaaaaa						1986

<210> 104

<211> 1333

<212> DNA

<213> Homo sapiens

<400> 104

gaattcggca	cgagcccagg	agtgcagtgg	tatgatcata	gttcaccgta	gcctcaaact	60
cgtgggctca	agtgatcctc	cagccttaac	ctcccgaata	gcctggccta	taggtgcacg	120
ccacacacct	gactgctcag	tatgtaaaatt	tttactatgc	ctaagggtga	ccacctttta	180
atatgttttag	gagccatttg	tatttccttt	tgtttcccat	attgttttgt	tcctatccat	240
ttttctacta	tatcgttgat	atgttggtta	tttggttaggg	atatgaaccc	tttgacagta	300
atgagttgca	aatattttct	ttccaatttg	tcactgtctc	tttgcttatg	atggccttgt	360
catgagtttt	aaaaaatttt	tatgtagtct	gaataccagt	tttttttagtg	gtttctggat	420
tttgagtcac	aattagaatg	twtttctcaa	tccagagcaa	tagagtaatt	cacctaaatt	480
ctacatctaa	attttgaacc	tctgaagcat	attctggcat	aagatataag	ttatggatct	540
aacctaat	tttccgcagg	tgattaaacc	agttgttcca	atattattta	ttgaactggt	600
tgttttttcc	tgacgagttt	gagargctac	attgatctta	tcttagaato	cgatcatatg	660
atttagctgt	gtatctgctt	ctgtttctct	gtatctgttt	ctatttcatt	gctctattta	720
gtcatgcact	artaccacat	tgttttaatt	acctaggctt	tagttttaat	ctagtgcatt	780
ggtcctccct	cattcctccc	ctgcccacct	tttttttttt	taacagtttt	tctaactggt	840
ccttattttt	cccatatgrg	ctttaaaaaa	ttcttaacat	atagagcata	ctaaaactgt	900
ccaactcaag	ttctctccca	agggttgcac	ttttaaccac	ttattttgtc	actgttcttt	960
tgatactttm	cctgataaaag	atacactttt	tactactttt	aaattattac	agtgttctat	1020
ttggcagtgc	ccaaacaggt	gatggcagat	agaggcagga	tgcaatgcct	gtgtggaaag	1080
aatgtcatct	cagtgtctct	attttaagat	agtctctagg	aatgatttaa	ggactgttct	1140
catgtaaaat	ccctatttct	ttttttatct	cattacgaat	tatttgccca	aaagtggat	1200
atctgtcaaa	gattcataag	acaagaggga	gagaccctta	aataagtact	aaacttgtaa	1260
aatcaatatg	tggataaaaag	tgcaagtaca	agaagttact	ttggaaaaaa	aaaaaaaaaa	1320
aaaaaaaaact	cga					1333

<210> 105

<211> 944

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (889)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (896)

<223> n equals a,t,g, or c

<400> 105

gaaaaagtac	aagcccctct	caaattgggtc	aagtttcaaa	tattagaccc	acccatggca	60
aagacagatt	ttagtataat	actcctaaaa	ctacactgtc	tttttttttt	ttctgtcata	120
agtgtgcatt	gtgctcagtc	atttatttca	gtgacccaaa	cagagcccag	tccagctggt	180
tgtattttcc	ctgcagtggg	aagtggacta	gggcatgtg	actaagaaag	ccagcctggg	240

61

ggctgtcttt	tcacctacag	atgtttta	at	gtgcttaaca	ttatccaata	ctagcaaccg	300
agatagtcta	aataccacag	caggatctga	ttagcttttt	cagatcactg	cctttatttg		360
ctgtttgcaa	aaaagcttaa	tccagtgcga	gagatcaggc	ttcctgctga	gccctggggt		420
agtttctctc	attctttgtg	ttcacagtgg	caggcgtag	tgagcagatt	cctcctcctc		480
ctaaattaaa	gctgtaaagt	agtaactgta	gtagcaagg	ataaagagaa	ggaagaaaac		540
ccaagggaaa	aaagaagact	gtctattcat	accaagtagt	ttccttgata	tacacaaaag		600
aaagagtttc	taatatgaat	tcataaatac	tgacctcagt	gtctcttcta	ctcagtgcac		660
agctattaag	ttttattagg	tttcagttgt	aactactttg	tgtggatata	tgttacgttt		720
ttcatattta	tcctactcaa	tcaatctcag	ttttaccaga	agaattacat	ttattagcca		780
taacagtggc	ccttctctta	ttcttttcag	ggctgatata	ttttttattc	atgagatttc		840
aaaaagaact	atcaccacca	ctaacaaaaa	aaaaaaaaaa	aaaaaaaaa	cggcncctct		900
agaggatccc	tcgagggg	caagcttacg	cgtgcatggg	acgt			944

<210> 106

<211> 1172

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (904)

<223> n equals a,t,g, or c

<400> 106

ggcggggcga	ggactccagc	gtgcccaggt	ctggcatcct	gcacttgctg	ccctctgaca	60
cctgggaaga	tggccggccc	gtggaccttc	acccttctct	gtggtttgct	ggcagccacc	120
ttgatccaag	ccaccctcag	tcccactgca	gttctcatcc	tcggcccaaa	agtcacaaa	180
gaaaagctga	cacaggagct	gaaggaccac	aacgccacca	gcacccctga	gcagctgccg	240
ctgctcagtg	ccatgcggga	aaagccagcc	ggagcatccc	tgtgctgggc	agcctggtga	300
acaccgtcct	gaagcacrtc	atctggctga	aggtcatcac	agytaacatc	ctccagctgc	360
aggtgaagcc	ctcggccaat	gamcaggagc	tgctagtcaa	gatccccctg	gacatggtgg	420
ctggattcaa	cacgcccctg	gtcaagacca	tcgtggagtt	ccacatgacg	actgaggccc	480
aagccaccat	ccgcatggac	accagtgcga	gtggccccac	ccgcctggtc	ctcagtgcct	540
gtgccaccag	ccatgggagc	ctgcgcaccc	aactgctgca	taagctctcc	ttcctggtga	600
acgccttagc	taagcaggct	atgaacctcc	tagtgccatc	catgccaaag	tggcccaact	660
gatcgtgctg	gaagtgtttc	cctccagtga	agccctccgc	cctttgttca	ccctgggcat	720
cgaagccagc	tcggaagctc	agttttacac	caaagggtgac	caacttatac	tcaacttgaa	780
taacatcagc	tctgatcgga	tccagctgat	gaactctggg	attggctggt	tccaacctga	840
tgttctgaaa	aacatcatca	ctgaratcat	ccactccatc	ctgctgccga	accagaatgg	900
caanttaaga	ctgggggtccc	agtgtcattg	gtgaaggcct	tgggattcga	ggcagctgag	960
tcctcactga	ccaaggatgc	ccttgtgctt	actccagcct	ccttgtggaa	acccasctct	1020
cctgtctccc	agtgaagact	tggatggcag	ccatcaggga	argctgggtc	ccagctggga	1080
rtatgggtgt	gagctctata	gaccatccct	ctctgcaatc	aataaacact	tgccctgtgaa	1140
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aa			1172

<210> 107

<211> 427

<212> DNA

<213> Homo sapiens

<400> 107

ccacgcgtcc	ggtgggctca	ctgttgggct	ccagcctagt	ggcactgctg	tccttgccc	60
ggggctggct	gcactgcccc	aaggactttg	ggaacatcaa	caattgccgg	atggacctct	120
acttcttcc	gctggctggc	attcaggccg	tcacggctct	cctatttgtc	tggatcgctg	180
gacgctatga	gagggcgtcc	cagggccag	cctccacag	ccgtttcagc	agggacagg	240
gctgaacagg	ccctattcca	gcccccttgc	ttcactctac	cggacagacg	gcagcagctc	300
cagctctggt	ttccttctcg	gtttattctg	ttagaatgaa	atggttccca	taaataagg	360
gcatgagccc	ttcctcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	420
aaaaaaa						427

<210> 108
 <211> 1708
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (85)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (254)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (256)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (423)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (424)
 <223> n equals a,t,g, or c

<400> 108
 ctctgtgcgaa ttctggcagag ctctggggcca atatggcagc gccagcaac aagacagagc 60
 tggcctggag tccgcggctg gccngtgag taggtgattg tctgacaagc agaggcataga 120
 gctgggtcca ggccacccta ctggcccagag gcctctgtag ggctgggga ggcacctgcg 180
 gggccgccct cacaggaacc tccatctctc aggtccctcg ccgctccct cggggcctcc 240
 actgcagcgc actnncata gctctgaaca gtccctgggt cccagcccac cggaaccccg 300
 gcagaggccc accaaggctc tgggtgccct tgaggacctg tttgggcagg cgcctgggtg 360
 ggaacgggac aaggcgagct tcctgcagac ggtgcagaaa tttgcggasa cagcgtgcgt 420
 aannggggcc acattgactt catctacctg gccctgcgca agatgcggga gtatggtgtc 480
 gagcgggacc tggctgtgta caaccagctg ctcaacatct tccccaaagg ggtcttcccg 540
 cctcgcaaca tcatccagcg catcttcgtc cactaccctc ggcagcagga gtgtgggatt 600
 gctgtcctgg agcagatgga gaaccacggt gtgatgcccc acaaggagac ggagttcctg 660
 ctgattcaga tctttggacg caaaagctac cccatgctca agttggtgcg cctgaagctg 720
 tggttccctc gattcatgaa cgtcaacccc tcccagtgcc cccgggacct gcccaggac 780
 cctgtggagc tggccatggt tggcctgcgg cacatggagc ctgaccttag tgccagggtc 840
 accatctacc aggttccttt gcccagagac tcaacaggtg cagcagatcc ccccagccc 900
 cacatcgtag gaatccagag tcccgatcag caggccgccc tggcccggca caatccagcc 960
 cggcctgtct ttgttgaggg ccccttctcc ctgtggctcc gcaacaagtg tgtgtattac 1020
 cacatcctca gagctgactt gctgcccccg gaggagaggg aagtgggaaga gacgccggag 1080
 gagtgggaacc tctactacc gatgcagctg gacctggagt atgtgaggag tggctgggac 1140
 aactacagagt ttgacatcaa tgaagtggag gaaggccctg tcttcgccat gtgcatggcg 1200
 ggtgctcatg accaggcgac gatggctaag tggatccagg gcctgcagga gaccaaccca 1260
 accctggccc agatccccgt ggtcttccgc ctgcgccggg ccacccggga gctccagaca 1320
 tcctctgcag ggctggagga gccgcccctg cccgaggacc accaggaaga agacgacaac 1380
 ctgcagcgac agcagcagg ccagagctag tctgagccgg cgcgagggca crggctgtgg 1440
 cccgaggagg cgggtggactg aaggcatgag atgccctttg agtgtacagc aaatcaatgt 1500
 tttcctgctt ggggctctct tccctcatct ctagcagtat ggcatccct ccccaggatc 1560
 tcgggctgcc agcgatgggc aggcgagacc cctccagaat ctgcaggcgc ctctggttct 1620

63

ccgaattcaa ataaaaagg gcgggagcgc tggtggtgt gcgcaaaaaa aaaaaaaaaa 1680
 aaaaaaaaaa aaaaaaagg gcggccgc 1708

<210> 109

<211> 1487

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (78)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (948)

<223> n equals a,t,g, or c

<400> 109

ccgctgctga taactatggc atccccggg cctgcaggaa ttcggcacgg agctacggcg 60
 ccgcctggct cctgctgnca cctgcaggct cgtcgcggtt ggagcccacc caagacatca 120
 gcatcagcga ccagctgggg ggccaggacg tgcccgtgtt ccggaacctg tccctgctgg 180
 tgggtgggtgt cggcgccgtg ttctcactgc tattccacct gggcaccctg gagaggcgcc 240
 ggccgcatgc ggagagacca ggcgagcaca cccccctgtt gggccctgcc acggccagc 300
 ccctgctgct ctggaagcac tggctccggg agcsggcttt ctaccagggtg ggcatactgt 360
 acatgaccac caggctcatc gtgaacctgt cccagacctt catggccatg tacctcacct 420
 actcgtccca cctgccccaa aagttcatcg cgaccattcc cctggtgatg tacctcagcg 480
 gcttcttctc ctcttctc atgaagccca tcaacaagtg cattgggagg aacatgacct 540
 acttctcagg cctcctgggtg atcctggcct ttgcccgtg ggtggcgctg gcggaggggac 600
 tgggtgtggc cgtgtacgca gcggctgtgc tgctgggtgc tggctgtgcc accatcctcg 660
 tcacctcgct ggccatgacg gccgacctca tcggtcccca cacgaacagc ggagckttcg 720
 tgtacggctc catgagcttc ttggataagg tggccaatgg gctggcagtc atggccatcc 780
 agagcctgca cccttgcccc tcagagctct gctgcagggc ctgctgagc ttttaccact 840
 gggcgatggt ggctgtgacg ggccggcgtg gcgtggccgc tgccctgtgt ctctgtagcc 900
 tcctgctgtg gccgaccgc ctgcgacgct gatgagacct gcacgcantg gctcacagca 960
 gcacgatttg tgacagcccg agcgaggaga caccgaacac ccagtgaagg tgaggggatc 1020
 agcacggcgc ggccaccac gcaccacgc gctggaatga gactcagcca caaggagggtg 1080
 cgaagctctg acccaggcca cagtgcggat gcacctgag gatgtcacgc tcagtggagag 1140
 acaccagaca cagaagggtg cgctgtgatc ccacttctat gaaatgtcca ggacagacca 1200
 atccacagaa tcaggggagag gattcgtggg tgccgggact ggggaggggg acctgggggt 1260
 gactagggtg cataatgggg acagggtgc cttctgggtg atgagaatgt tctggaatca 1320
 gatgggatgg ctgcacggcg tggtaagggt actgaacgcc acctcactgt aagacggtag 1380
 atttgtatt ttaccacaat aaacaaaaca aaacaaaacc aaaaaaaaaa aaaaaaaaaa 1440
 aaaaaaagg aattcgatat caagcttctc gataccgtcg acctcga 1487

<210> 110

<211> 1525

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (78)

<223> n equals a,t,g, or c

<400> 110

ccgctgctga taactatggc atccccggg cctgcaggaa ttcggcacgg agctacggcg 60
 ccgcctggct cctgctgnca cctgcaggct cgtcgcggtt ggagcccacc caagacatca 120
 gcatcagcga ccagctgggg ggccaggacg tgcccgtgtt ccggaacctg tccctgctgg 180

64

tggtgggtgt	cggegcctgt	ttctcactgc	tattccacct	gggcacccgg	gagaggcgcc	240
ggccgcatgc	ggasgagcca	ggcgagcaca	ccccctgtt	ggccctgcc	acggccagc	300
ccctgctgct	ctggaagcac	tggtccggg	agcsggcttt	ctaccagggtg	ggcatactgt	360
acatgaccac	caggctcatc	gtgaacctgt	cccagacct	catggccatg	tacctacct	420
actcgtcca	cctgcccag	aagtccatcg	cgaccattcc	cctgggtgatg	tacctacg	480
gcttcttgtc	ctccttctc	atgaagccca	tcaacaagt	cattgggagg	aacatgacct	540
acttctcagg	cctcctgggtg	atcctggcct	ttgccgcctg	ggtggcgctg	gcggaggag	600
tggtgtggc	cgtgtacgca	gcggctgtgc	tgctgggtgc	tggtgtgccc	accatcctcg	660
tcacctcgct	ggccatgacg	gccgacctca	tcggtcccca	cacgaacagc	ggactktcgt	720
gtacggctcc	atgagcttct	tggataaggt	ggccaatggg	ctggcagtc	tggccatcca	780
gagcctgcac	ccttgcccct	cagagctctg	ctgcagggcc	tgcgtgagct	tttaccactg	840
ggcgatgggtg	gctgtgacgg	gcggcggtggg	cgtggccgct	gccctgtgtc	tctgtagcct	900
cctgctgtgg	ccgaccggcc	tgcgacgctg	ggaccgtgat	gcccggccct	gactcctgac	960
agcctcctgc	acctgtgcaa	gggaactgtg	gggacgcacg	aggatgcccc	ccarggcctt	1020
gggaaaagc	ccccactgcc	cctcactctt	ctctggaccc	ccaccctcca	tcctcaccca	1080
gctcccgggg	gtgggtcg	gtgagggcag	cagggtgccc	cgccaggag	ttgcaaggac	1140
ccctggggtt	ttgagggtgt	cccattctca	actctaattc	atcccagccc	tctggaggat	1200
ttgggggtgcc	cctctcggca	gggaacagga	agttaggaatc	ccagaagggt	ctgggggaac	1260
cctaaccctg	agctcagtc	agttcacccc	tcacctccag	cctgggggtc	tccagact	1320
gccaggggcc	cctcaggacg	gctggagcct	ggaggagaca	gccacggggt	ggtgggctgg	1380
gcctggaccc	caccgtgggtg	ggcagcagg	ctgcccggca	ggcttgggtg	actctgctgg	1440
cagcaaataa	agagatgacg	gcaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1500
aaaaaaaaaa	aaaccaccg	tccgc				1525

<210> 111

<211> 552

<212> DNA

<213> Homo sapiens

<400> 111

ccacgcgtcc	ggtcagaatg	ccttggaaaa	gagctgtagt	tctcctaata	ttatggttta	60
tagggcaggc	catgtggctg	gctcctgcct	atgttctaga	gtttcaagga	aagaacacct	120
ttctgtttat	ttggttagct	ggtttgttct	ttcttcttat	caattgttcc	atcctgattc	180
aaattatttc	ccattacaaa	gaagaacccc	tgacagagag	aatcaaatat	gactagtgt	240
tgttccacac	cctctgctac	tgtgttacat	tctgattgtc	ttgtatggac	cagaagagag	300
ctttgggaca	ttttttctga	acattctaa	cattctagtg	aaagtccca	tgttccaaca	360
gaacttaaaa	gcaatgtttg	ccttatatat	aaaagggaca	caataattga	ggtccacctt	420
ctaggaaatc	ctaggactcg	tttatttggg	acatgggtgg	aataaagggtc	acatattgga	480
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	540
aaaaaaaaaa	aa					552

<210> 112

<211> 925

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (444)

<223> n equals a,t,g, or c

<400> 112

ctgcaggaat	tcggcacgag	cggaaccggg	gccggctgct	gtgcatgctg	gcgctgacct	60
tcatgttcat	ggtgctggag	gtgggtggtga	gccgggtgac	ctcgtcgctg	gcgatgctct	120
ccgactcctt	ccacatgctg	tcggacgtgc	tggcgctggt	ggtggcgctg	gtggccgagc	180
gcttcgccc	gcggacccac	gccaccaga	agaacacgtt	cggctggatc	cgagccgagg	240
taatgggggc	tctggtgaac	gccatcttcc	tgaactggcct	ctgtttcgcc	atcctgctgg	300
aggccatcga	gcgcttcac	gagccgcacg	agatgcagca	gccgctgggtg	gtccttgggg	360
tcggcgctgg	cgggctgctg	gtcaacgtgc	tggggctctg	cctcttccac	catcacagcg	420

65

gcttcagcca	ggactccggc	cacngccact	cgcacggggg	tcacggccac	ggccacggcc	480
tccccaaggg	gcctcgcggt	aagagcacc	gccccgggag	cagcgacatc	aacgtggccc	540
cgggcgagca	gggtcccgac	caggaggaga	ccaacaccct	ggtggccaat	accagcaact	600
ccaacgggct	gaaattggac	cccgcagacc	cagaaaaccc	cagaagtggg	gatacagtgg	660
aagtacaagt	gaatggaaat	cttgtcagag	aacctgacca	tatggaactg	gaagaagata	720
gggctggaca	acttaacatg	cgtggagttt	ttctgcatgt	ccttggagat	gccttggggt	780
cagtgattgt	agtagtaaat	gccttagtct	tttacttttc	ttggaaaggt	tggtctgaag	840
gggatttttg	tgtgaatcca	tgtttccttg	accctgcaa	agcatttgta	gaaatattaa	900
tagtactcat	gcatacagttt	atgag				925

<210> 113

<211> 1340

<212> DNA

<213> Homo sapiens

<400> 113

ggcacgagaa	agaaaggcga	gagaaaaatc	aaggcaccaa	atttagattg	gaggtctcag	60
aggagcagtg	ttttccctcc	ttcgtaacag	ttgaacaact	tccagatgta	gctagctgca	120
ccccctgtaa	agatgcaggc	tctttacaat	gaagacacat	cttctgatgt	tccttctctc	180
ctgtatggcc	agatgcacag	gaatagtgcc	caaaagacct	cagcctgctt	tccctttaag	240
gggaaggaga	agaaaaaact	cctttttatt	tttactttct	ttcagcattg	aatttttggt	300
gtgtgtatgg	tgacttctgt	ttttgggaaa	cgggaagaag	ccagcagcat	gctgaattgt	360
cctgacaggc	tccgctgggc	tcttgccgag	gttagcagtg	ctttttttgt	atttaaacca	420
tctcccgggc	agtgtaaaaa	gtttgcaggt	gcggacattc	tgtctgactg	gtctcggcag	480
tgctctataa	ccctgttggt	tttcttgata	aaacacagcc	ccacccttta	ataaagcaaa	540
gattgctatg	aaaccagaga	gtctattcat	tactgtggag	taactagagc	agtctgtagt	600
gactagacat	acggcaatta	ggaagtcag	gagttgggat	ttttgtctta	attttggctg	660
ctcaaagtgc	cccctgtagg	atattctttt	ttcgggaatt	gtttccaaac	ttgcctgtct	720
ttatctatgg	tgaaactcaa	gccgcttttt	aaggcaagcc	tgcaaacc	agtatcaaca	780
tgggctcctg	aaggcacagg	gagcagattc	acagttctga	ccagtgttag	ggtccccacg	840
agggccaccc	atttgaactc	aagggtggca	gactctggcc	ccagcacttg	ccgtgggttc	900
aggatggcca	gcggtgacac	agggctatgg	aaccctgggt	cttcactctc	tcccatatcc	960
tttgtttcac	cttctttttg	ccatatttta	ttgtgcttca	gatagaaatt	ttattttataa	1020
gataaaaaagt	agctctgagg	ctgggcacgg	tggtctcatgc	ctgtgggtccc	agcacttttg	1080
gaggccgagg	tggttggttc	acgagctcag	cagatcaaga	ccatcctggc	caatatgggtg	1140
aaacctgtc	tctgctaaaa	atacaaaaat	tggtctggcg	tggtggcggg	tgctgtagt	1200
cccagctact	cgggaggctg	aggcgggaga	atcgattgga	cccaggaggc	ggaggttgca	1260
gtgagcctag	atggcaccac	tgcgctccag	cctgggtgac	agagggagac	tgctcaaaa	1320
aaaaaaaaaa	aaaaaaaaaa					1340

<210> 114

<211> 813

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (338)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (384)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (389)

<223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (799)
 <223> n equals a,t,g, or c

<400> 114
 ctgcaggaat tcggcacgag aaagaaaggc gagagaaaaa tcaaggcacc aaatttagat 60
 tggaggtctc agaggagcag tgttttccct ccttcgtaac agttgaacaa cttccagatg 120
 tagctagctg caccctctgt aaagatgcag gctctttaca atgaagacac atcttctgat 180
 gttccttctc tcctgtatgg ccagatgcac aggaatagtg cccaaaagac ctcagcctgc 240
 tttcccttta agggggaagg agaagaaaaa actccttttt atttttactt tctttcagca 300
 ttgaattttt gttgtgtgta tggtgacttc tggttttngg gaaacggaag aagccagcag 360
 catgctgaat tgcctgaca ggcntccgnt ggctcttgcc gaggttagca gtgctttttt 420
 tgwattttaa ccatctcccg ggcagtgtaa aaagtttgca ggtgcggaca ttctgtctga 480
 ctggtctcgg cagtgtctta taacctgtt gtgtttcttg ataaaacaca gccccacct 540
 ttaataaagc aaagattgct atgaaaccag agagtctatt cattactgtg gactaactag 600
 agcagtctgt agtgactaga catacgcaa ttaggaagtc atggagttgg gatttttgct 660
 ttaatttttg ctgctcaaag tgccccctgt aggatattct ttttctggga attgtttcca 720
 aacttgctg tctttatcta tggtgaaact caagccgctt tttaaggcaa gcctgcaaac 780
 ccaagtatca acatggggnc ctgaagggac agg 813

<210> 115
 <211> 1681
 <212> DNA
 <213> Homo sapiens

<400> 115
 cgatggcccc gcggccgctc tagaaagtcc cgtttttttt tttttttttt tttttttttt 60
 ttttagagta cgttctgcat tttatttytg caggcaacac tttgctcacc agcaagaaca 120
 cagcccragg aagggaacca ataacctttc aaaacscaaa ctgctkcctg cggtaggggc 180
 ccagggtcct ccacggagag gacaggcatc ttcctttccc accaggaagg agtcagcccc 240
 gagcctctgc tatgtgcaag gcggtgtgca agcaccggct gcggctcttt gctgtctctt 300
 ctttctcttt ggggctgggc tgggtgtgct ttctgtgct gatgctttgg cctgtgaggc 360
 tgagcttggc ayctcgaccc gttcaattac agcaacgaag aagccactgc tragygtggt 420
 ctcaggggar gcccggaggc agtgctggc acccggaac gtgctcaggc ctcgggtggg 480
 ccaggcaggc agggcgaggc cttagcctgaa ggcgccccgg ttctgctgca gcgcactctc 540
 caccacgtct tcattctcct cctggcagag ggagcacgtg gagtagacga gccgctgcag 600
 ggaagggaaa gtgagcgcgt ggcacagggc tcgctgctgg aaccctgcca gggcatgcag 660
 acgcaccggg ctagggtgtsc ctgccccggg mtctccagc tgtctgctcg gcatacccca 720
 gccactgcag gaaggatcca gcaggayrta gtggacctca ygrtagcgyg gatcyraggg 780
 ggagaccgcc aggaagtcct cctcagccag ytcacagcar gagacgccag cccrggccag 840
 cagcgtggcc atggatgcca gccgcttggc atccaggctc aaggcaaaga tcttcccttg 900
 gttcttcaga agagcagcca agtgactggc cttattgcct ggggcgggc aggcactgat 960
 gacatgggag cctggcgggg ggtccagcag catggctggg agacagctgg ccctgtcctg 1020
 cagaatgagg tgtccggccc ggtacagtgg gtgttcagc agatctgtct gggcgggaaa 1080
 caccagcagc tccggcatca aggggtccag gagaaaatgc ttccccttga gggctcgtaa 1140
 gtcactgagg ctggaagccc gaccctgata ggagaaacct tgtctcttga aataatcaac 1200
 tacatcatcg gagcaggtct tgagagtgtt cacacgcaca aatcgaggca gctgggaggc 1260
 tggaccaggc cttgatccca cttccaacag gtcctcattc cggctcacac cccgatgaac 1320
 cttgagccga gccaaactcag ctttgagcct cgcttggcgc cggcccaaca ggccttcca 1380
 tcggccccc cccctcgaa agccctttcc caacaacaac tcataacta gcaccttggc 1440
 cagggtgcggc cgcagcttct tctccgcagc gaggagggcg gcgctggcga tcacagctc 1500
 cagcacggcg gagtagcgct gcgtttcgca caccagcgcg tacagctgct tcacgttctg 1560
 gaagttgctg gagtacacca accccttgat agagcctggc ggctctccac gccggccaac 1620
 acgcctgcag ctgcagcata cagccccatg ttccgtcgcg ctttacggct ttgtggcaaa 1680
 a 1681

<210> 116

<211> 2052
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (2045)
<223> n equals a,t,g, or c

<400> 116
tttgcttttc aaatgctccc aaggtctcag atgaagcggg gaaaaaagat tcagagttgg 60
ataagcactt ggaatcacgg gttgaagaga ttatggagaa gtctggcgag gaaggaatgc 120
ctgatcttgc ccatgtcatg cgcattctgt ctgcagaaaa tatcccaaat ttgcctcctg 180
ggggaggtct tgctggcaas cgtaattgta ttgaagctgt ttatagtaga ctgaatccac 240
atagagaaaag tgatgggggt gctggagatc tagaagacct atggtagcct taaaaacctt 300
ctaaaatgct tttrattctg aaaattgggg gaaaaaactt ttaatcaca ttttcttcaa 360
tacaagggga aaatattctt gcggattccc aacgttttgt gatatgagca gaaaatcatt 420
agcatttccc atcatttgtt catatttgtg ttttctgaca gttgccactt gtagcattgc 480
ctgtactaca gtattttttg ccaacctcag gcatactcgt tacatctgta ttgaactttc 540
ggccctagaa accagtggag ttatttcacc acaaatcaac aatgtgcctg aggtgcattg 600
gaaatatagt tagctatact ctgaaaatac attatgtttt ttttctttaa acaaaacaca 660
caacatgtaa gcatgtaaga gtaagaatt gtatgatatg ttcccttttt cagttcacca 720
agttggaagc cttttgcagc tctgtggcct ggaatttcat ttgagcaatt tctataggat 780
atgtatttat tattgattgt tatttaawtt ttttcccaat tttacctgta ttaccaaact 840
gggttctcca ataattgtcca aattgtaatg ttgccttgct tcaagataaa gtgtatttgg 900
gaataatatt ataaacctt acaatttta tgcatgtatc tactgcatcc ttcaactctc 960
actagaaaat cttttgaaac caaatggatt aatttatggc tatttataat ttgctttgac 1020
atctcactgt tggaaatttt ttaaagatga gatttgcctt tataatgtaa attgtgattt 1080
ttgttttaca tgtgggtttc tatagtttta attttttcag cttttaagat acgagttttg 1140
tgtaatttgg tatttttaatt catttatgtt attttaaaag ctcagaatat cacattgaaa 1200
ttactataaa tacatttaaa attatctatt ttagatctaa ggaaatacta cagagatatt 1260
ttcatgggtt cagtaacttt tcattttata acattgggca cggtagagag tgattgtcac 1320
ataaggtact tgaagattta ttagtttaat tctattttta cagtaacctt gaattcttct 1380
gagttttgca tgtattaaat tcaattaatg ctgaacatga agagtaaagt atttatctga 1440
aagaagtttc tgggttagga gaagtaatga atgtatccat ttgtacatgg ttacatggtt 1500
gtggatgctt tgtaaacatt ttcctgtatg tttaaattgt gtttcagcag gatgtaattg 1560
cccttggtg tagttaaaat gagtcatcat ctggctcttt gtgaaatgga attcatggta 1620
ttttctgtaa cgttttcctg aagctgtttc tggagagcca cacatttaaa tacagacagc 1680
tttcctgatc atttgattta ttgtgcacct gatttttggg ctaaaaggaa ttattgccac 1740
aatatatattt atttattctt tagatttttag ccttgtaagt taaagtgcct tacatgatga 1800
tgtgaaaagc tgtttgtccc ttactgggtt ttgggggggt gttaaaagat aggggaatgaa 1860
gaatgcaaaa tggtttatcg ttcaactgt ccactctgat ccaaccctgt actgatagta 1920
cttccagta tgatattgtg atgtttcata caatgcagtg aacataacca acttgttacc 1980
taaataaaga attgataaaa acagtgtgac atattaaaaa aaaggggggc ccggtaccca 2040
attcnccta ta 2052

<210> 117
<211> 539
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (528)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (529)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (531)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (532)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (537)

<223> n equals a,t,g, or c

<400> 117

gagatacatt	ccatgaatac	ctagttttatt	gagagttttt	agcatgaagg	actgtcgaat	60
tttgtcaaaag	gctttttctg	catctattga	gataatcatg	tggtttttgt	ctttggttct	120
gtttatgtga	tggaactatgt	ttattgattt	gcataatgtt	aaccagcctt	gcattctcagg	180
gatgaagcca	actcgatcgt	tgtggataag	ctttttgatg	tgctgctgga	tttggtttgc	240
caatatttta	ttgaggattt	ttgcatcagt	gttcttcagg	gatattgggtc	taaaattctc	300
ttttttttgt	tgtgtctctg	ccaggctttg	gtatcaggat	gatgctggcc	tcataaatga	360
gttagggagg	attccctctt	tctattgatc	agaatagttt	cagaaggaat	ggtagcagct	420
cttctttgta	cctctggtag	aatttggtg	kgaatctatc	ttgkcctgga	atatttttgg	480
ggttggaact	caaaaaaaaa	aaaaaaaaaa	tcaaaaaaaaa	aaaaaaanna	nnaaaaanaa	539

<210> 118

<211> 882

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (117)

<223> n equals a,t,g, or c

<400> 118

gaattcggca	cgagcagacc	tgggctcgag	accataactg	tttggcttta	acagtacgtg	60
ggcggccgga	atccgggagt	ccggtgaccc	gggctgtggt	ctagcataaa	ggcggancca	120
gaagaagggg	cgggggatgg	gagaagcctc	cccacctgcc	cccgaaggc	ggcatctgct	180
ggctcctgctg	ctgctcctct	ctaccctggt	gatccccctc	gctgcagctc	ctatccatga	240
tgctgacgcc	caagagagct	ccttgggtct	cacaggcctc	cagagcctac	tccaaggctt	300
cagccgactt	ttcctgaaag	taacctgctt	cggggcatag	acagcttatt	ctctgcccc	360
atggacttcc	ggggcctccc	tgggaactac	cacaaagagg	agaaccagga	gcaccagctg	420
gggaacaaca	ccctctccag	ccacytccag	atcgacaaga	tgaccgacaa	caagacagga	480
gaggtgctga	tctccgagaa	tgtggtggca	tccattcaac	cagcggaggg	gagcttcgag	540
ggtgatttga	aggtacccag	gatggaggag	aaggaggccc	tggtacccat	ccagaaggcc	600
acggacagct	tccacacaga	actccatccc	cgggtggcct	tctggatcat	taagctgcca	660
cggcggaggt	cccaccagga	tgccctggag	ggcgccact	ggctcagcga	gaagcgacac	720
cgctgcagg	ccatccggga	tggactccgc	aaggggaccc	acaaggacgt	cctagaagag	780
gggaccgaga	gtcctccca	ctccaggctg	tcccccgaa	agaccactt	actgtacatc	840
ctcaggccct	ctcggcagct	gtaggggtgg	ggaccgggga	gc		882

<210> 119

<211> 1193

<212> DNA

<213> Homo sapiens

<400> 119

acactatata	agttacgcct	gcagggttacc	gggccggtaa	ttcccgggtc	gtacccacgc	60
gtccggtaat	gtcaaaggaa	aagtaattct	gtcaatgctg	gttgtctcaa	ctgtgatcat	120
tgtgttttgg	gaatttatca	acagcacaga	aggctctttc	ttgtggatat	atcactcaaa	180
aaaccagaa	gttgatgaca	gcagtgtca	gaaggctgg	tggtttctga	gctggtttaa	240
caatgggatc	cacaattatc	aacaagggga	agaagacata	gacaaagaaa	aaggaagaga	300
ggagacaaaa	ggaaggaaaa	tgacacaaca	gagcttcggc	tatgggactg	gtttaatcca	360
aacttgaagg	aatccgaata	actaaactgg	actctggttt	tctgactcag	tccttctaga	420
agacctggac	tgagagatca	tgccgttaag	gagtgtgtaa	caggcggacc	acctgttggg	480
actgsgagat	tctcaagggg	aaggactggg	tctcattttc	cccattctcag	cgcttagcag	540
gatgacctgg	tatagagcag	ggaactggga	aatgtgggtc	aggggatcag	acactccagt	600
tgggtctttt	atataaatta	aatggcaaaa	ggctccatac	ccttctcctt	ctttcctacc	660
ctccacttta	tctgcaaaa	gggaatgatg	ataacaccca	cttcatagaa	tggtcatgaa	720
gatcaaatga	gagaataaaa	gtcaagcact	tagcctctgg	tgacaataaa	gtattaaata	780
agtataccta	ttcctccttt	tcctttttta	aaaataatat	taccaaatgt	ccagcttata	840
cacatttaca	agacttagct	agtgggctat	gttagagcta	ctaaaagatc	tttgacaagc	900
taaaactaag	atgcaatgaa	tgagggtgaa	cgaacaagag	agttttaagt	tcagaaatgg	960
ttacagaagt	ataagacagc	tgtgtgggtg	ttttttgggt	tttggtttct	ggtttacaat	1020
ctcgtcatte	aacaaagatg	ggagttttat	agaactaaaa	gcmccatgta	agctactaaa	1080
aacaacaaca	aaaaaggctc	atcattttct	agtctgaatt	gacaaaaatg	ccaatgcaaa	1140
taaaaatgat	tactttttat	tttaaaaaaa	aaaaaaaaaa	aaaaaaactc	gta	1193

<210> 120

<211> 1338

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (519)

<223> n equals a,t,g, or c

<400> 120

ggcacgaggg	tgaggcccag	gtagcgtttg	caatccagcc	ccaccgtcac	ctcttttctt	60
ggacttctag	ttttcctcac	ccctattgcc	ttcatccttt	tacctccgat	cctgtggagg	120
gaatgagctg	gagccttggt	gcacaatttg	tgaggggctc	tttatctcca	tggtattcaa	180
actcctcatt	ctgctcatag	ggacctgggc	actttttttc	cgcaagcgga	gagctgacat	240
gccacgggtg	tttgtgtttc	gtgccctttt	gttggtcctc	atctttctct	tttgtggttt	300
ccctattggc	ttttttacgg	ggcccgcat	ttggactctc	gggaaccgga	attaccaagg	360
gattgtgcaa	tatgcagtct	ccccttggtg	aatgccctcc	tccttccatc	cattactggc	420
catccgtccc	tgtctggagct	cagggagctt	gcagcccaat	gttccacgct	gcaggttggg	480
cccgtcccca	accgaatggg	gaaatccccg	cttccagcnt	gggacacctg	agtatccagc	540
gagcagcatt	gggtgtccta	gaaaattact	acaaagattt	caccatctat	aacccaaacc	600
tcctaacagc	ctccaaattc	cgagcagcca	agcatatggc	cgggctgaaa	gtctacaatg	660
tagatggccc	cagtaacaat	gccactggcc	agtcccgggc	catgattgct	gcagctgctc	720
ggcgagggga	ctcaagccac	aacgagttgt	attatgaaga	ggccgaacat	gaacggcgag	780
taaaagaagc	gaaagcaagg	ctggtggttg	cagtgggaaga	ggccttcac	cacattcagc	840
gtctccaggg	tgaggagcag	cagaaagccc	caggggaggt	gatggacctt	agggaggccg	900
cccaggccat	tttccctcc	atggccaggg	ctctccagaa	gtacctgcgc	atcaccgggc	960
agcagaacta	ccacagcatg	gagagcatcc	tgcaagcacc	tggccttctg	catcaccaac	1020
ggcatgaccc	ccaaggcctt	cctagaacgg	tacctcagtg	cgggccccac	cctgcaatat	1080
gacaaggacc	gctggctctc	tacacagtgg	aggcttgta	gtgatgaggc	tttgactaat	1140
ggattacggg	atggaattgt	gttcgtcctt	aagtgttgg	acttcagcct	cgtagtcaat	1200
gtgaagaaaa	ttccattcat	catactctct	gaagagttca	tagaccccaa	atctcacaaa	1260
ttgtccttc	gcttacagtc	tgagacatcc	gtttaaaagt	tctatatattg	tggctttatt	1320
aaaaaaaaaa	aaaaaaaaaa					1338

<210> 121

<211> 1183

<212> DNA

<213> Homo sapiens

<400> 121

tgcaggaatt	cggcacgagc	tggtctgcagg	gtctctgggg	agagaagggg	cctcgggttc	60
acaggatggg	gctgccagt	tcctggggcc	ctcctgccct	ctgggttcta	gggtgctgcg	120
ccctgtcct	ctcgtgtg	gcgctgtgca	cagcctgccg	cagcccagg	acgctgtagc	180
ccccaggaag	agggcgcgga	ggcagcgggc	gaggctgcag	ggcagtgcga	cggcggcgga	240
agcgtcccta	ctgaggcgga	cccacctctg	cttccctcag	caagtccggac	accagactgc	300
acgagctgca	ccggggcccg	cgcagcagca	gggccctgcg	gcctgccagy	atggatctcc	360
tgcgcccaca	ctggctggag	gtgtccagg	acatcaccgg	accgcaggca	gccccctctg	420
ccttcccaca	ccaggagctg	ccccgggctc	tgccggcagc	tgcagccacc	gcagggtgcg	480
tggcctcgag	gccacctatt	ccaacgtggg	gctggcgggc	cttcccgggg	tcagcctggc	540
ggccagccct	gtggtggccg	agtatgccc	cgtccagaag	cgcaaaggga	cccattcgag	600
tccccaagag	ccacagcagg	ggaagactga	ggtgaccccg	gccgctcagg	tggacgtcct	660
gtactccagg	gtctgcaagc	ctaaaaggag	ggacccagga	cccaccacag	acccgctgga	720
ccccaagggc	cagggagcga	ttctggccct	ggcgggtgac	ctggcctacc	agacctccc	780
gctcagggcc	ctggatgtg	acagcgcccc	cctggaaaac	gtgtatgaga	gcattccgga	840
gctgggggac	cctgctggca	ggagcagcac	gtgcggggct	gggacgccc	ctgcttccag	900
ctgccccagc	ctaggagggg	gctggagacc	cctccctgcc	tcctgccct	gaacactcaa	960
ggacctgtgc	tccttcctcc	agagtgaggc	ccgtcccccg	ccccgccccg	cctcacagct	1020
gacagcgcca	gtcccaggtc	cccgggccc	cagcccgtag	ggtccgtgag	gtcctggccg	1080
ctctgacagc	cgcgccctcc	ccgggctcca	gagaaggccc	gcgtctaaat	aaagcgccag	1140
cgcaggatga	aagcgaaaaa	aaaaaaaaaa	aaagggcggc	cgc		1183

<210> 122

<211> 615

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (18)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (20)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (584)

<223> n equals a,t,g, or c

<400> 122

cctgtatata	aaattggn	ctatggctcc	gtacaatgaa	gaaatgcaa	gatagttaag	60
aaagactcgg	ccttcaagga	gcctaaatgt	gtagaaaagg	actaaggcaa	aacaataact	120
tttttgagct	cttgccatgt	gtgaagcact	ttatacacct	gtaaggtagg	taacgttgtt	180
cttattaaac	atgaagaaaa	tgagactttg	tgagaagcaa	tacagtatag	aagttaagaa	240
tatggactct	aaagctagat	ttcagagggt	tgaagttagt	ctgctactta	ctggctgtgt	300
gactttgagc	agattactta	acctgtctgt	gcctatgttt	actttttattg	ttgtaaaaaag	360
atatgcaaca	taaaatattc	catttcaacc	gtttttacgt	gtatacttca	ctgacattag	420
ttgcatccac	tatgttgtgc	aaacgtagg	tcgctatgaa	gattaaatga	gttaattcat	480
ataaagccct	cagaagagtg	tctggcacat	ggtgagtatt	ggctgtactg	tggctgatgt	540
cattgttaga	gagcttttagt	gatttgctta	agacagaaa	gtanactggg	gtgcggtggt	600
ctcacgccct	ggtta					615

<210> 123
<211> 587
<212> DNA
<213> Homo sapiens

<400> 123
cccacgcgtc cgcttggaac ctgattctcc tgaccgtctt taccctgtcc atggcctacc 60
tcaactgggat gctgtccagc tactacaaca ccacctccgt gctgctgtgc ctgggcatca 120
cggcccttgt ctgcctctca gtcaccgtct tcagcttcca gaccaagttc gacttcacct 180
cctgccaggg cgtgctcttc gtgcttctca tgactctttt cttcagcggg ctcactcctgg 240
ccatcctcct acccttccaa tatgtgccct ggctccatgc agtttatgca gcaactgggag 300
cgggtgtatt tacattgttc ctggcacttg acaccagtt gctgatgggt aaccgacgcc 360
actcgctgag ccctgaggag tatatttttg gaggccctaa catctaccta gacatcatct 420
atatcttcac cttcttctg cagctttttg gcactaaccg agaatgagga gccctccctg 480
ccccaccgtc ctccagagaa tgcgcccctc ctgggtccct gtccctcccc tgcgtcctg 540
cgagaccaga tataaaacta gctgccaacc caaaaaaaaa aaaaaaa 587

<210> 124
<211> 1379
<212> DNA
<213> Homo sapiens

<400> 124
gggcccagca gcagcggcac ctggagaagc agcacctgcg aattcagcat ctgcaaagcc 60
agtttggcct cctggaccac aagcacctag accatgaggt ggccaagcct gcccgaagaa 120
agaggctgcc cgagatggcc cagccagttg acccggctca caatgtcagc cgctgcacc 180
ggctgcccag ggattgccag gagctgttcc aggttgggga gaggcagagt ggactatttg 240
aaatccagcc tcaggggtct ccgccatttt tgggtgaactg caagatgacc tcagatggag 300
gctggacagt aattcagagg cgccacgatg gctcagtggg cttcaaccgg ccctgggaag 360
cctacaaggc ggggtttggg gatccccacg gcgagttctg gctgggtctg gagaaggtgc 420
atagcatcat gggggaccgc aacagccgcg tggcctgca gctgcgggac tgggatggca 480
acgccagatt gctgcagttc tccgtgcacc tgggtggcga ggacacggc tatagcctgc 540
agctcactgc acccgtggcc ggccagctgg gcgccaccac cgtcccacc agcggcctct 600
ccgtaccctt ctccacttgg gaccaggatc acgacctccg cagggaacaag aactgcgcca 660
agagcctctc tggaagctgg tggtttggca cctgcagcca ttccaacctt caacgggcca 720
gtacttccgg ctccatccca cagcagcggc agaagcttaa gaagggaatc ttctggaaga 780
cctgcggggc gctactaccc gctgcaggcc accaccatgt tgatccagcc catggcagca 840
gaggcagcct cctagcgtcc tggctgggccc tgggtcccagg ccacgaaaag acggtgactc 900
ttggtctctc ccgaggatgt ggccgttccc tgcctgggca ggggctccaa ggaggggcca 960
tctggaaact tgtggacaga gaagaagacc acgactggag aagccccctt tctgagtgc 1020
ggggggctgc atgcgttggc tcctgagatc gaggctgcag gatatgctca gactctagag 1080
gcgtggacca aggggcatgg agcttcaact cttgctggcc agggagtgg ggactcagag 1140
ggaccacttg gggccagcca gactggcctc aatggcggac tcagtcacat tgactgacgg 1200
ggaccagggc ttgtgtgggt cgagagcgcc ctcatggtgc tgggtgctgt gtgtgtaggt 1260
cccctgggga cacaagcagg cgccaatggt atctgggcgg agctcacaga gttcttggaa 1320
taaaagcaac ctcagaacac ttaaaaaaaaa aaaaaaaaa aaaaaaaaa 1379

<210> 125
<211> 1268
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1184)
<223> n equals a,t,g, or c

<220>
<221> SITE

<222> (1240)

<223> n equals a,t,g, or c

<400> 125

```

agggttgatg gggttatggtc aggagtgccc gctggggccc ccacctcctc aggaaggcgg      60
gtgaggttg  tgtgagactg acggtgcctc ctcatgtccc cttggagcgc cccaccccac      120
atctcccggc ctccgggtcct tgcctggccc agcatgagag gtgcttcata ggaacggagg      180
gaggacatgt cgggacagct cgatgctcgg cctgctgctg ctctgcaccc ccagggcctg      240
gctcacccctc tctggacctg tctgcttcca aggaaggggac cctctgaggt ccacagagg      300
ccaccccagc tgtgggtcgt gagcatctct gtcttgacag gacagcatcg tggccgagct      360
ggaccgagag atgagcagag cgtggacgtg accaacacca ccttcctgct catggccgcc      420
tccatctatc tccacgacca gaacccggat gccgccctgc gtgcgctgca ccagggggac      480
agcctggagt ggtgagtggc ctccctgctc tgggccagcc cagggaggca agtgccccct      540
gccacatctc caggctgcgc acggcctcgc tggctgtcgt catgggagca gagaaagggt      600
gtgctgaaat gaggccttg  cctgctgtcc aggtccagc tccctgccc agtgtgggag      660
gcactcccat ctgcgcacca ggctgcggat ccaaggacac ggtgccagg ctgcaaccct      720
ctgttcccaa gggcagagca gaaagcggct ttgtctctgc tcggtttctg tgtccccacc      780
ccccacgaa  ccttctgtgt ctccggccctg ggccagtcct ctcaggcctc cccgggcccc      840
ccataccgyc cctcctccag ggccctctgg ggttgggggt ctgaagccct gcaaggtttg      900
tgccccctc  caccctagga tgtgactccg ggccatgtcc agggcactgg tcacagaaag      960
tgtgtcagtt cttccccgtg agctgtccct gcagtgcctg ccttccactg tgagttgcaa     1020
gctgggcatt tcatggtcgc tgtggatctg ctcccatccc acctccatcc acagagggct     1080
tagaattgca gggcgagcca ggcattggtga catgcaccta tgtttccagc tacttgggag     1140
gcggaagtca ggagtatccc ttgagtctgg gaggtggagg ctgncagtga gccgtgatgg     1200
tgccactgca ctccagcctg ggtggcagag ccagaccctn actcacacac aaaaaaaaaa     1260
aaaaaaaaa                                     1268

```

<210> 126

<211> 1311

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1036)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1112)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1168)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1223)

<223> n equals a,t,g, or c

<400> 126

```

gaaaaaagaa agcaatatgg aaaccgaact aaggagattt taaactgaga tataagatgc      60
tttcaattat tccaatgac aggctattta tcaatttaaat atttttaagc aacttcctcc     120
catcagtgtc ctgggaacca gctgggcaga tgtggtacac ccatgtcaga taccocagtg     180
gcaggctcct gtcactgtag cacttggtcc ctccatccct cccagccttc ctagctcctt     240
gctcctggaa acctcccccc atcaatctct gacatttcag aggaaatact gtttgtcacc     300
tcttaaggaa tctgggagga cggcctgtga gatatggcgt cagttacagc ctcttaaaga     360

```

73

gtcaatagcc	cctgcagagg	ccagaacact	ggaacaaatg	taaggaaggt	atagttttta	420
aagatttttg	acttgaatta	aataggattg	gttactttct	gcccctcccg	aggggtggact	480
gtgcacagaa	gagacctctt	caccgggttt	gctgctcttt	ttcgcaactgt	gagttgggggt	540
tctaacagtc	agcgttggtc	cataacaaaa	tggaatcct	ttctttcccc	tcctgttaat	600
gcccctgtc	tgtgcagtga	ctgtgcaacc	agcacctttt	gtggtcgaat	cagccagcag	660
aagtgccct	cgtgttcctg	gattctctct	tctgtgggtc	catttctttg	agtcctgggt	720
tctcgccctg	aatggctcaa	cagggggaaa	ggcagacagc	ttcttcgtgc	cagaaacatt	780
ttttttttt	tttgaaatar	tgagccaaga	ttgcgccact	gcattccatc	ctcagcaaca	840
garcaagact	ccaactcawa	acaaaacaaa	agattgargt	wattgtggca	acacctgcct	900
ttttttctaa	gctgcaattc	tctactgttt	tcaagaaaaa	tacaagttag	cctattttaca	960
gaatgttttg	aattgactcc	tgctctctgg	ttaaaactcc	tcttgagata	attgatagct	1020
gaaaaggtag	gatgntctc	tcaaacttga	cttccatcta	aatcaacgct	gagttgatta	1080
acttagatat	caagaaaaat	tgctcatta	gnttaccct	gaggagatgc	ctatgaaggt	1140
acatcctttt	tacaattaat	aagacagntt	tcacatgaag	aaacaatttg	aaatatttaa	1200
taagaaaatg	gggtgaaggc	aancattacg	gttgggaaaa	gaccatgcaa	gcctttatag	1260
aggataacga	tttatatatt	cactattaat	ttggccgggt	aataggaacc	t	1311

<210> 127

<211> 1249

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1217)

<223> n equals a,t,g, or c

<400> 127

ggccaggcgg	gtctcaaaact	cctcgtctca	ggtgatctgc	ttgcctcggc	ctcccaaagt	60
gctgggatta	caggcgtgag	cactgcgccc	agcctgagtt	tcatttttta	agtcacatag	120
cagtagtcct	tatttcagtg	ctagaccctt	tgaaatgcga	tgaaagctat	atggaccctt	180
cgctttgtta	tataacatat	gcacacatac	ccagaatttt	gcacatatgt	tcagagattc	240
ctagacctgc	agacctgcct	ctgtgtgtcc	caatttaaga	acctctgttc	tttcttcagt	300
actggatttg	cccaattttg	tgttattttg	ggacttaatt	tgtccctctt	tgggacattt	360
ccttatttat	tgccctcttc	agagagtaga	tgtagaaaat	aaagagagga	aacctagatt	420
acttaatttt	aatttaacat	tttctataga	tagcatacca	cgccaagtgt	gctctgtctt	480
gatccccctc	tttctagcat	ctgccagaca	ttgtagagtt	tcscascag	ttgtagggtt	540
gagctgcagc	cagtcatttc	ttttattctt	taaaagtaca	tagatttgct	tttttagggc	600
tttactgaaa	gtaaaatata	ctgacattta	aactgacaga	tgtaggaggt	aaaaaataga	660
gttctgaaac	atwtgaattt	atgtgacagc	tgaagtcacg	agatgaggka	tgtatgtccc	720
ccagggaggw	tgcaaaaaga	agaaaagggt	actggaaaca	gcatgtcagt	ggtgccagct	780
gagggctgga	ggcagccagg	agagttggga	gcctgggtgc	tgggtggaga	gaggttaaca	840
gggaakacat	gggaagtatt	gtgaaggctg	gtgtgagcag	gggactactc	cagccctggt	900
ggaacataga	gccatttggt	agattgacaa	tgcaagtaca	gctgtatata	ataaatgtgt	960
tgaaaggagg	aaggtgagga	ttttcttggt	gggagtttat	gctgttattt	aacatatttt	1020
gcttccaaag	gggttaagat	gttttaccta	aatggargtt	tctaggtcag	tgctatacaa	1080
tattttcta	ctgtgtttta	tagtgtgagc	tacatatgta	attttaaaa	tttcaagtag	1140
ccacataata	aaggaaacag	gtgaaattta	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1200
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa		1249

<210> 128

<211> 1660

<212> DNA

<213> Homo sapiens

<400> 128

ccgggtcgac	ccacgcgtcc	ggcccccgga	aggcgacatg	ggctccgctc	cctggggccc	60
ggctcctgctg	ctggcgctcg	ggctgcgcgg	cctccaggcg	ggggcccgca	gcggaccccg	120
gcttccagga	gcgcttcttc	cagcagcgtc	tggaccactt	caacttcgag	cgcttcggca	180

74

acaagacctt	ccctcagcgc	ttcctgggtg	cggacaggtt	ctgggtccgg	ggcgaggggc	240
ccatcttctt	ctacactggg	aacgagggcg	acgtgtgggc	cttcgccaac	aactcggcct	300
tcgtcgcgga	ctggcgggcg	agcggggggc	tctactggtc	ttcgcggagc	accgctacta	360
cgggaagtcg	ctgccgttcg	gtgcgcagtc	cacgcagcgc	gggcacacgg	agctgtgtac	420
ggtggagcag	gccctggccg	acttcgcaga	gctgtctccg	gcgctacgac	gcgacctcgg	480
ggcccaggat	gcccccgcca	tcgccttcgg	tggaagttat	ggggggatgc	tcagtgccta	540
cctgaggatg	aagtatcccc	acctgggtgg	ggggggcgctg	gcggccagcg	cgcccgttct	600
atctgtggca	ggcctcggcg	actccaacca	gttcttccgg	gacgtcacgg	cggactttga	660
gggccagagt	cccaaagtga	cccagggtgt	gcgggaagcg	ttccgacaga	tcaaggactt	720
gttcctacag	ggagcctacg	acacgttccg	ctgggagttc	ggcacctgcc	agccgctgtc	780
agacgagaag	gacctgacct	agctcttcat	gttcgcccgg	aatgccttca	ccgtgtctgg	840
catgatggac	tacccttacc	ccactgactt	cctgggtccc	ctccctgcca	accccgtcaa	900
ggtgggctgt	gatcggctgc	tgagttaggc	ccagaggatc	acggggctgc	gagcactggc	960
agggtgtgtc	tacaacgcct	cgggctccga	gcactgtctac	gacatctacc	ggctctacca	1020
cagctgtgct	gacccactg	gctgcggcac	cggccccgac	gccagggcct	gggactacca	1080
ggcctgcacc	gagatcaacc	tgaccttcgc	cagcaacaat	gtgaccgata	tgttcccgga	1140
cctgcccttc	actgacgagc	tccgccagcg	gtactgcctg	gacacctggg	gcgtgtggcc	1200
cgggcccgac	tggtgtctga	ccagcttctg	gggggggtgat	ctcagagccg	ccagcaacat	1260
catcttctcc	aacgggaacc	tggaccctcg	gggagggggc	gggattcggg	ggaacctgag	1320
tgccctcagtc	atcgccgtca	ccatccaggg	gggagcgcac	cacctcgacc	tcagagcctc	1380
ccaccagaa	gatcctgctt	ccgtggttga	ggcgcggaag	ctggaggcca	ccatcatcgg	1440
cgagtgggta	aaggcagcca	ggcgtgagca	gcagccagct	ctgcgtgggg	ggcccagact	1500
cagcctctga	gcacaggact	ggaggggtct	caaggctcct	catggagtgg	gggcttact	1560
caagcagctg	gcggcagagg	gaaggggctg	aataaacgcc	tgagggcctg	gccatgtaaa	1620
aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa			1660

<210> 129

<211> 2075

<212> DNA

<213> Homo sapiens

<400> 129

ccacgcgtcc	gtggggccga	gcgccgctgg	gtaggcgga	gtagccgcag	atggcggcgg	60
ctatgccctt	gctctgctcg	tcctgttgct	cctggggccc	ggcggtggt	gccttgca	120
acccccacgc	gacagcctgc	gggaggaact	tgatcatcacc	ccgtgcctt	ccggggacgt	180
agccgccaca	ttccagttcc	gcacgcgctg	ggattcggag	cttcagcggg	aaggagtgtc	240
ccattacagg	ctctttccca	aagccctggg	gcagctgctc	tccaagtatt	ctctacggga	300
gctgcacctg	tcattcacac	aaggcttttg	gaggaccgga	tactgggggc	cacccttcc	360
gcaggcccca	tcagacactg	accactactt	tctgcgctat	gctgtgctgc	cgcgggaggt	420
ggtctgcacc	gaaaacctca	ccccctggaa	gaagctcttg	ccctgtagtt	ccaaggcagg	480
cctctctgtg	ctgctgaagg	cagatcgctt	gttccacacc	agctaccact	cccaggcagt	540
gcataatccgc	cctgtttgca	gaaatgcacg	ctgtactagc	atctcctggg	agctgaggca	600
gaccctgtca	gttgtatttg	atgccttcat	cacggggcag	ggaaagaaag	actggtccct	660
cttcgggatg	ttctcccga	ccctcacgga	gccctgcccc	ctggcttcag	agagccgagt	720
ctatgtggac	atcaccacct	acaaccagga	caacgagaca	ttagagggtg	accaccccc	780
gaccactaca	tatcaggacg	tcattcctag	cactcggaag	acctatgcca	tctatgactt	840
gcttgacacc	gccatgatca	acaactctcg	aaacctcaac	atccagctca	agtggagag	900
acccccagag	aatgaggccc	ccccagtgcc	cttcctgcat	gcccagcggg	acgtgagtgg	960
ctatgggctg	cagaaggggg	agctgagcac	actgctgtac	aacacccacc	cataccgggc	1020
cttcccgggtg	ctgctgctgg	acaccgtacc	ctggtatctg	cggtgtatg	tgacacccct	1080
caccatcacc	tccaagggca	aggagaacaa	accaagttac	atccactacc	agcctgcccc	1140
ggaccggctg	caacccccacc	tcctggagat	gctgattcag	ctgccggcca	actcagtcac	1200
caagggttcc	atccagtttg	agcggggcgt	gctgaagctg	accgagtaca	caccagatcc	1260
taacctggtg	ttctatgtca	gccatctgt	cctcagcgc	cttgtgcccc	gcatggtagc	1320
agccaagcca	gtggactggg	aagagagtcc	cctcttcaac	agcctgttcc	cagtctctga	1380
tggtcttaac	tactttgtgc	ggctctacac	ggagccgctg	ctggtgaacc	tgccgacacc	1440
ggacttcagc	atgccctaca	acgtgatctg	cctcacgtgc	actgtgggtg	ccgtgtgcta	1500
cggctccttc	tacaatctcc	tcacccgaac	ctttccacat	cgaggagccc	cgcacagggtg	1560
gcctggccaa	gcggctggcc	aaccttatcc	ggcgcgcccc	agtgcccccc	ccactctgat	1620

75

tcttgccctt	tccagcagct	gcagctgccg	tttctctctg	gggaggggag	cccaagggct	1680
gtttctgcca	cttgctctcc	tcagagttgg	cttttgaacc	aaagtgccct	ggaccaggtc	1740
agggcctaca	gctgtgtgt	ccagtacagg	agccacgagc	caaagtgtgc	atttgaattt	1800
gaattaactt	agaaattcat	ttcctcacct	gtagtggcca	cctctatatt	gaggtgctca	1860
ataagcaaaa	gtggtcgggtg	gctgctgtat	tggacagcac	agaaaaagat	ttccatcacc	1920
acagaaagggt	cggctggcag	cactggccaa	ggtgatgggg	tgtgctacac	agtgtatgtc	1980
actgtgtagt	ggatggagtt	tactgtttgt	ggaataaaaa	cggctgtttc	cgtgggtaaa	2040
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaa			2075

<210> 130

<211> 56

<212> PRT

<213> Homo sapiens

<400> 130

Met	Ala	Lys	Thr	Asp	Phe	Ser	Ile	Ile	Leu	Leu	Lys	Leu	His	Cys	Leu
1				5					10					15	

Phe	Phe	Phe	Ser	Val	Ile	Ser	Val	His	Cys	Ala	Gln	Ser	Phe	Ile	Ser
			20					25					30		

Val	Thr	Gln	Thr	Glu	Pro	Ser	Pro	Ala	Val	Cys	Ile	Phe	Pro	Ala	Val
		35					40					45			

Gly	Ser	Gly	Leu	Gly	Pro	Cys	Asp
	50					55	

<210> 131

<211> 42

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (3)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (42)

<223> Xaa equals stop translation

<400> 131

Met	Ala	Xaa	Leu	Asp	Asn	Cys	Leu	Met	Leu	Leu	Ile	Thr	Ser	Gly	Thr
1				5					10					15	

Trp	Leu	Gly	Ser	Val	Ala	Arg	Lys	Thr	Trp	Gln	Ala	Ile	Cys	Asp	Ser
			20					25					30		

Gly	Ser	Ser	Gly	Cys	Ala	Leu	Ile	Arg	Xaa
	35						40		

<210> 132

<211> 415

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

76

<222> (415)

<223> Xaa equals stop translation

<400> 132

Met Asn Pro Thr Leu Gly Leu Ala Ile Phe Leu Ala Val Leu Leu Thr
 1 5 10 15

Val Lys Gly Leu Leu Lys Pro Ser Phe Ser Pro Arg Asn Tyr Lys Ala
 20 25 30

Leu Ser Glu Val Gln Gly Trp Lys Gln Arg Met Ala Ala Lys Glu Leu
 35 40 45

Ala Arg Gln Asn Met Asp Leu Gly Phe Lys Leu Leu Lys Lys Leu Ala
 50 55 60

Phe Tyr Asn Pro Gly Arg Asn Ile Phe Leu Ser Pro Leu Ser Ile Ser
 65 70 75 80

Thr Ala Phe Ser Met Leu Cys Leu Gly Ala Gln Asp Ser Thr Leu Asp
 85 90 95

Glu Ile Lys Gln Gly Phe Asn Phe Arg Lys Met Pro Glu Lys Asp Leu
 100 105 110

His Glu Gly Phe His Tyr Ile Ile His Glu Leu Thr Gln Lys Thr Gln
 115 120 125

Asp Leu Lys Leu Ser Ile Gly Asn Thr Leu Phe Ile Asp Gln Arg Leu
 130 135 140

Gln Pro Gln Arg Lys Phe Leu Glu Asp Ala Lys Asn Phe Tyr Ser Ala
 145 150 155 160

Glu Thr Ile Leu Thr Asn Phe Gln Asn Leu Glu Met Ala Gln Lys Gln
 165 170 175

Ile Asn Asp Phe Ile Ser Gln Lys Thr His Gly Lys Ile Asn Asn Leu
 180 185 190

Ile Glu Asn Ile Asp Pro Gly Thr Val Met Leu Leu Ala Asn Tyr Ile
 195 200 205

Phe Phe Arg Ala Arg Trp Lys His Glu Phe Asp Pro Asn Val Thr Lys
 210 215 220

Glu Glu Asp Phe Phe Leu Glu Lys Asn Ser Ser Val Lys Val Pro Met
 225 230 235 240

Met Phe Arg Ser Gly Ile Tyr Gln Val Gly Tyr Asp Asp Lys Leu Ser
 245 250 255

Cys Thr Ile Leu Glu Ile Pro Tyr Gln Lys Asn Ile Thr Ala Ile Phe
 260 265 270

Ile Leu Pro Asp Glu Gly Lys Leu Lys His Leu Glu Lys Gly Leu Gln
 275 280 285

Val Asp Thr Phe Ser Arg Trp Lys Thr Leu Leu Ser Arg Arg Val Val

77

290 295 300
 Asp Val Ser Val Pro Arg Leu His Met Thr Gly Thr Phe Asp Leu Lys
 305 310 315 320
 Lys Thr Leu Ser Tyr Ile Gly Val Ser Lys Ile Phe Glu Glu His Gly
 325 330 335
 Asp Leu Thr Lys Ile Ala Pro His Arg Ser Leu Lys Val Gly Glu Ala
 340 345 350
 Val His Lys Ala Glu Leu Lys Met Asp Glu Arg Gly Thr Glu Gly Ala
 355 360 365
 Ala Gly Thr Gly Ala Gln Thr Leu Pro Met Glu Thr Pro Leu Val Val
 370 375 380
 Lys Ile Asp Lys Pro Tyr Leu Leu Leu Ile Tyr Ser Glu Lys Ile Pro
 385 390 395 400
 Ser Val Leu Phe Leu Gly Lys Ile Val Asn Pro Ile Gly Lys Xaa
 405 410 415
 <210> 133
 <211> 45
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (45)
 <223> Xaa equals stop translation
 <400> 133
 Met Gly Gln Gln Ser Cys Trp Met Gly Leu Gly Cys Trp Leu Ser Leu
 1 5 10 15
 Ser Gly Leu Ser Gly Val Val Arg Ala Ser Pro Arg Ser Pro Arg Pro
 20 25 30
 Arg Arg Gly Ala Ala Cys Gly Glu Thr Leu Met Pro Xaa
 35 40 45
 <210> 134
 <211> 197
 <212> PRT
 <213> Homo sapiens
 <400> 134
 Met Ala Gly Pro Trp Thr Phe Thr Leu Leu Cys Gly Leu Leu Ala Ala
 1 5 10 15
 Thr Leu Ile Gln Ala Thr Leu Ser Pro Thr Ala Val Leu Ile Leu Gly
 20 25 30
 Pro Lys Val Ile Lys Glu Lys Leu Thr Gln Glu Leu Lys Asp His Asn
 35 40 45
 Ala Thr Ser Ile Leu Gln Gln Leu Pro Leu Leu Ser Ala Met Arg Glu

[illegible]

79

<220>

<221> SITE

<222> (41)

<223> Xaa equals stop translation

<400> 136

Met Leu Lys Arg Val Ile Leu Leu Val Glu Met Phe Ile His Phe Leu
1 5 10 15

Ile Tyr Ala Lys Ser Phe Tyr His Lys Ser Trp Glu Gln Leu Ser Phe
20 25 30

Thr His Tyr Leu Leu Gln Ile Ser Xaa
35 40

<210> 137

<211> 85

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (85)

<223> Xaa equals stop translation

<400> 137

Met Pro Ile Leu Val Phe Ser Ile Cys Leu Gln Cys Thr Leu Phe Arg
1 5 10 15

Ser Glu Ala Ile Ile Phe Gln Glu Glu Arg Asn His Gln Val Thr Leu
20 25 30

Leu Lys Ala Val Lys Thr Lys Phe Gln Ser Gly Thr Gly Leu Arg Xaa
35 40 45

Pro Val Leu Glu Tyr Ala Lys Ser Ile Gln Ile Ile Ser Lys Tyr Thr
50 55 60

Cys Gly Thr Val Leu Pro Val Phe Lys Met Arg Arg Tyr Tyr Val Gly
65 70 75 80

Gln Lys Cys Gln Xaa
85

<210> 138

<211> 201

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (144)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (149)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (160)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (173)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (177)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (189)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (201)
 <223> Xaa equals stop translation

<400> 138
 Met Phe Phe Leu Leu Cys Leu Val Ala Leu Glu Ile Lys Gly Phe Thr
 1 5 10 15
 Phe Ser Ala Arg Gly Ala Arg Asp Arg Phe Leu Asn Lys Ser Gly Pro
 20 25 30
 Gln Pro Gly Lys Lys Met Lys Thr Thr His Cys Lys Gln Pro Leu Phe
 35 40 45
 Ser Lys Pro Gly Gln Val Arg Gly Ala Leu Arg Lys Ala Arg Gly Arg
 50 55 60
 Gln Glu Glu Arg Glu Ala Val Gly Met Trp Gly Gly Arg Gly His Ser
 65 70 75 80
 Tyr Pro Glu Tyr Ile Lys Thr Ser Glu Val Thr Glu Val Arg Asp Ser
 85 90 95
 Pro Lys His Pro Gln Val Gln Pro Phe Leu Thr Thr Arg Val Thr Cys
 100 105 110
 Arg Val Pro Gly His Leu Gln Val Leu Glu Ala Leu Cys Gly Ala Trp
 115 120 125
 Gly Ser Met Phe Lys His Ala Leu Val Val Val Gln Val Pro Arg Xaa
 130 135 140
 Arg Gly Arg Ala Xaa Leu Gly Ser Glu Trp Gln Val Gly Gln Leu Xaa

145		150		81		155		160
Leu Ile Leu Leu His Gly Thr Gln His Trp Ala Ala Xaa Leu Val Pro		165		170				175
Xaa Leu Pro Gln Glu Ser Ile Leu Pro Ala Gln Ser Xaa Arg Val Thr		180		185				190
Asn Thr Pro Gly Thr Glu Glu Thr Xaa		195		200				
<210> 139								
<211> 325								
<212> PRT								
<213> Homo sapiens								
<400> 139								
Met Gly Ser Gln Val Ser Ser Met Leu Lys Leu Ala Leu Gln Asn Cys		5		10				15
Cys Pro Gln Leu Trp Gln Arg His Ser Ala Arg Asp Arg Gln Cys Ala		20		25				30
Arg Val Leu Ala Asp Glu Arg Ser Pro Gln Pro Gly Ala Ser Pro Gln		35		40			45	
Glu Asp Ile Ala Asn Phe Gln Val Leu Val Lys Ile Leu Pro Val Met		50		55			60	
Val Thr Leu Val Pro Tyr Trp Met Val Tyr Phe Gln Met Gln Ser Thr		65		70			75	80
Tyr Val Leu Gln Gly Leu His Leu His Ile Pro Asn Ile Phe Pro Ala			85			90		95
Asn Pro Ala Asn Ile Ser Val Ala Leu Arg Ala Gln Gly Ser Ser Tyr		100		105				110
Thr Ile Pro Glu Ala Trp Leu Leu Leu Ala Asn Val Val Val Val Leu		115		120			125	
Ile Leu Val Pro Leu Lys Asp Arg Leu Ile Asp Pro Leu Leu Leu Arg		130		135			140	
Cys Lys Leu Leu Pro Ser Ala Leu Gln Lys Met Ala Leu Gly Met Phe			150			155		160
Phe Gly Phe Thr Ser Val Ile Val Ala Gly Val Leu Glu Met Glu Arg			165			170		175
Leu His Tyr Ile His His Asn Glu Thr Val Ser Gln Gln Ile Gly Glu		180			185			190
Val Leu Tyr Asn Ala Ala Pro Leu Ser Ile Trp Trp Gln Ile Pro Gln		195		200			205	
Tyr Leu Leu Ile Gly Ile Ser Glu Ile Phe Ala Ser Ile Pro Gly Leu		210		215			220	

82

Glu Phe Ala Tyr Ser Glu Ala Pro Arg Ser Met Gln Gly Ala Ile Met
 225 230 235 240

Gly Ile Phe Phe Cys Leu Ser Gly Val Gly Ser Leu Leu Gly Ser Ser
 245 250 255

Leu Val Ala Leu Leu Ser Leu Pro Gly Gly Trp Leu His Cys Pro Lys
 260 265 270

Asp Phe Gly Asn Ile Asn Asn Cys Arg Met Asp Leu Tyr Phe Phe Leu
 275 280 285

Leu Ala Gly Ile Gln Ala Val Thr Ala Leu Leu Phe Val Trp Ile Ala
 290 295 300

Gly Arg Tyr Glu Arg Ala Ser Gln Gly Pro Ala Ser His Ser Arg Phe
 305 310 315 320

Ser Arg Asp Arg Gly
 325

<210> 140

<211> 119

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (107)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (119)

<223> Xaa equals stop translation

<400> 140

Met Val Phe Val His Leu Tyr Leu Gly Asn Val Leu Ala Leu Leu Leu
 1 5 10 15

Phe Val His Tyr Ser Asn Gly Asp Glu Ser Ser Asp Pro Gly Pro Gln
 20 25 30

His Arg Ala Gln Gly Pro Gly Pro Glu Pro Thr Leu Gly Pro Leu Thr
 35 40 45

Arg Leu Glu Gly Ile Lys Val Gly His Glu Arg Lys Val Gln Leu Val
 50 55 60

Thr Asp Arg Asp His Phe Ile Arg Thr Leu Ser Leu Lys Pro Leu Leu
 65 70 75 80

Phe Glu Ile Pro Gly Phe Leu Thr Asp Glu Glu Cys Arg Leu Ile Ile
 85 90 95

His Leu Ala Gln Met Lys Gly Leu Gln Arg Xaa Arg Ser Cys Leu Leu
 100 105 110

Lys Ser Met Lys Arg Gln Xaa

115

<210> 141
<211> 48
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (8)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (19)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (48)
<223> Xaa equals stop translation

<400> 141
Met Lys Leu Thr Ile Phe Phe Xaa Phe Pro Gln Thr Ile Thr Gly Leu
1 5 10 15

Leu Gln Xaa Leu Met Ser Arg Gln Val Glu Asp Val Ala Phe Leu Pro
20 25 30

Leu Pro His Pro Val Phe Ser Phe Ser Phe Phe Phe Pro Leu Val Xaa
35 40 45

<210> 142
<211> 520
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (205)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (207)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (213)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (225)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (520)

<223> Xaa equals stop translation

<400> 142

Met Gln Gly Gly Gln Arg Pro His Leu Leu Leu Leu Leu Ala Val
 1 5 10 15

Cys Leu Gly Ala Gln Ser Arg Asn Gln Glu Glu Arg Leu Leu Ala Asp
 20 25 30

Leu Met Arg Asn Tyr Asp Pro His Leu Arg Pro Ala Glu Arg Asp Ser
 35 40 45

Asp Val Val Asn Val Ser Leu Lys Leu Thr Leu Thr Asn Leu Ile Ser
 50 55 60

Leu Asn Glu Arg Glu Glu Ala Leu Thr Thr Asn Val Trp Ile Glu Met
 65 70 75 80

Gln Trp Cys Asp Tyr Arg Leu Arg Trp Asp Pro Lys Asp Tyr Glu Gly
 85 90 95

Leu Trp Ile Leu Arg Val Pro Ser Thr Met Val Trp Arg Pro Asp Ile
 100 105 110

Val Leu Glu Asn Asn Val Asp Gly Val Phe Glu Val Ala Leu Tyr Cys
 115 120 125

Asn Val Leu Val Ser Pro Asp Gly Cys Ile Tyr Trp Leu Pro Pro Ala
 130 135 140

Ile Phe Arg Ser Ser Cys Ser Ile Ser Val Thr Tyr Phe Pro Phe Asp
 145 150 155 160

Trp Gln Asn Cys Ser Leu Ile Phe Gln Ser Gln Thr Tyr Ser Thr Ser
 165 170 175

Glu Ile Asn Leu Gln Leu Ser Gln Glu Asp Gly Gln Ala Ile Glu Trp
 180 185 190

Ile Phe Ile Asp Pro Glu Ala Phe Thr Glu Asn Gly Xaa Trp Xaa Ile
 195 200 205

Arg His Arg Pro Xaa Lys Met Leu Leu Asp Ser Val Ala Pro Ala Glu
 210 215 220

Xaa Ala Gly His Gln Lys Val Val Phe Tyr Leu Leu Ile Gln Arg Lys
 225 230 235 240

Pro Leu Phe Tyr Val Ile Asn Ile Ile Ala Pro Cys Val Leu Ile Ser
 245 250 255

Ser Val Ala Ile Leu Ile Tyr Phe Leu Pro Ala Lys Ala Gly Gly Gln
 260 265 270

Lys Cys Thr Val Ala Thr Asn Val Leu Leu Ala Gln Thr Val Phe Leu

Met Leu Leu Phe Ser Ser Arg Phe Ile Met Phe Leu Trp Pro Pro Val

1 5 86 10 15
 Ser Gly Val Cys Leu Ser Phe Ile Arg Asp Arg Ser Phe Leu Pro Met
 20 25 30
 Cys His Phe Ile Tyr Val Leu Ile Leu Cys Asn Ser Ile Ala Leu Xaa
 35 40 45

 <210> 144
 <211> 431
 <212> PRT
 <213> Homo sapiens

 <400> 144
 Met Ser Trp Val Gln Ala Thr Leu Leu Ala Arg Gly Leu Cys Arg Ala
 1 5 10 15
 Trp Gly Gly Thr Cys Gly Ala Ala Leu Thr Gly Thr Ser Ile Ser Gln
 20 25 30
 Val Pro Arg Arg Leu Pro Arg Gly Leu His Cys Ser Ala Ala Ala His
 35 40 45
 Ser Ser Glu Gln Ser Leu Val Pro Ser Pro Pro Glu Pro Arg Gln Arg
 50 55 60
 Pro Thr Lys Ala Leu Val Pro Phe Glu Asp Leu Phe Gly Gln Ala Pro
 65 70 75 80
 Gly Gly Glu Arg Asp Lys Ala Ser Phe Leu Gln Thr Val Gln Lys Phe
 85 90 95
 Ala Glu His Ser Val Arg Lys Arg Gly His Ile Asp Phe Ile Tyr Leu
 100 105 110
 Ala Leu Arg Lys Met Arg Glu Tyr Gly Val Glu Arg Asp Leu Ala Val
 115 120 125
 Tyr Asn Gln Leu Leu Asn Ile Phe Pro Lys Glu Val Phe Arg Pro Arg
 130 135 140
 Asn Ile Ile Gln Arg Ile Phe Val His Tyr Pro Arg Gln Gln Glu Cys
 145 150 155 160
 Gly Ile Ala Val Leu Glu Gln Met Glu Asn His Gly Val Met Pro Asn
 165 170 175
 Lys Glu Thr Glu Phe Leu Leu Ile Gln Ile Phe Gly Arg Lys Ser Tyr
 180 185 190
 Pro Met Leu Lys Leu Val Arg Leu Lys Leu Trp Phe Pro Arg Phe Met
 195 200 205
 Asn Val Asn Pro Phe Pro Val Pro Arg Asp Leu Pro Gln Asp Pro Val
 210 215 220

87

Glu Leu Ala Met Phe Gly Leu Arg His Met Glu Pro Asp Leu Ser Ala
 225 230 235 240

Arg Val Thr Ile Tyr Gln Val Pro Leu Pro Lys Asp Ser Thr Gly Ala
 245 250 255

Ala Asp Pro Pro Gln Pro His Ile Val Gly Ile Gln Ser Pro Asp Gln
 260 265 270

Gln Ala Ala Leu Ala Arg His Asn Pro Ala Arg Pro Val Phe Val Glu
 275 280 285

Gly Pro Phe Ser Leu Trp Leu Arg Asn Lys Cys Val Tyr Tyr His Ile
 290 295 300

Leu Arg Ala Asp Leu Leu Pro Pro Glu Glu Arg Glu Val Glu Glu Thr
 305 310 315 320

Pro Glu Glu Trp Asn Leu Tyr Tyr Pro Met Gln Leu Asp Leu Glu Tyr
 325 330 335

Val Arg Ser Gly Trp Asp Asn Tyr Glu Phe Asp Ile Asn Glu Val Glu
 340 345 350

Glu Gly Pro Val Phe Ala Met Cys Met Ala Gly Ala His Asp Gln Ala
 355 360 365

Thr Met Ala Lys Trp Ile Gln Gly Leu Gln Glu Thr Asn Pro Thr Leu
 370 375 380

Ala Gln Ile Pro Val Val Phe Arg Leu Ala Gly Ser Thr Arg Glu Leu
 385 390 395 400

Gln Thr Ser Ser Ala Gly Leu Glu Glu Pro Pro Leu Pro Glu Asp His
 405 410 415

Gln Glu Glu Asp Asp Asn Leu Gln Arg Gln Gln Gln Gly Gln Ser
 420 425 430

<210> 145

<211> 443

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (364)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (443)

<223> Xaa equals stop translation

<400> 145

Met Trp Phe Thr Tyr Leu Leu Leu Tyr Leu His Ser Val Arg Ala Tyr
 1 5 10 15

Ser Ser Arg Gly Ala Gly Cys Cys Cys Cys Trp Ala Arg Trp Arg Arg

88

20	25	30
Ala Val His Thr Ala Arg Gly Leu Arg Gly Arg Pro Arg Arg Gln Leu 35 40 45		
Leu Arg Pro Leu Arg Pro Ala Gln Gly Leu Ala Pro Gly Arg His Arg 50 55 60		
Leu Arg Pro Ala Val Leu Pro Leu His Leu Gln Pro Leu Pro Gly Leu 65 70 75 80		
Trp Gly Gly His Ala Glu Trp Ala Ala Leu Leu Tyr Tyr Gly Pro Phe 85 90 95		
Ile Val Ile Phe Gln Phe Gly Trp Ala Ser Thr Gln Ile Ser His Leu 100 105 110		
Ser Leu Ile Pro Glu Leu Val Thr Asn Asp His Glu Lys Val Glu Leu 115 120 125		
Thr Ala Leu Arg Tyr Ala Phe Thr Val Val Ala Asn Ile Thr Val Tyr 130 135 140		
Gly Ala Ala Trp Leu Leu Leu His Leu Gln Gly Ser Ser Arg Val Glu 145 150 155 160		
Pro Thr Gln Asp Ile Ser Ile Ser Asp Gln Leu Gly Gly Gln Asp Val 165 170 175		
Pro Val Phe Arg Asn Leu Ser Leu Leu Val Val Gly Val Gly Ala Val 180 185 190		
Phe Ser Leu Leu Phe His Leu Gly Thr Arg Glu Arg Arg Arg Pro His 195 200 205		
Ala Glu Glu Pro Gly Glu His Thr Pro Leu Leu Ala Pro Ala Thr Ala 210 215 220		
Gln Pro Leu Leu Leu Trp Lys His Trp Leu Arg Glu Pro Ala Phe Tyr 225 230 235 240		
Gln Val Gly Ile Leu Tyr Met Thr Thr Arg Leu Ile Val Asn Leu Ser 245 250 255		
Gln Thr Tyr Met Ala Met Tyr Leu Thr Tyr Ser Leu His Leu Pro Lys 260 265 270		
Lys Phe Ile Ala Thr Ile Pro Leu Val Met Tyr Leu Ser Gly Phe Leu 275 280 285		
Ser Ser Phe Leu Met Lys Pro Ile Asn Lys Cys Ile Gly Arg Asn Met 290 295 300		
Thr Tyr Phe Ser Gly Leu Leu Val Ile Leu Ala Phe Ala Ala Trp Val 305 310 315 320		
Ala Leu Ala Glu Gly Leu Gly Val Ala Val Tyr Ala Ala Ala Val Leu 325 330 335		

89

Leu Gly Ala Gly Cys Ala Thr Ile Leu Val Thr Ser Leu Ala Met Thr
 340 345 350

Ala Asp Leu Ile Gly Pro His Thr Asn Ser Gly Xaa Phe Val Tyr Gly
 355 360 365

Ser Met Ser Phe Leu Asp Lys Val Ala Asn Gly Leu Ala Val Met Ala
 370 375 380

Ile Gln Ser Leu His Pro Cys Pro Ser Glu Leu Cys Cys Arg Ala Cys
 385 390 395 400

Val Ser Phe Tyr His Trp Ala Met Val Ala Val Thr Gly Gly Val Gly
 405 410 415

Val Ala Ala Ala Leu Cys Leu Cys Ser Leu Leu Leu Trp Pro Thr Arg
 420 425 430

Leu Arg Arg Trp Asp Arg Asp Ala Arg Pro Xaa
 435 440

<210> 146

<211> 76

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (76)

<223> Xaa equals stop translation

<400> 146

Met Ser Arg Phe Ile Leu Asn His Leu Val Leu Ala Ile Pro Leu Arg
 1 5 10 15

Val Leu Val Val Leu Trp Ala Phe Val Leu Gly Leu Ser Arg Val Met
 20 25 30

Leu Gly Arg His Asn Val Thr Asp Val Ala Phe Gly Phe Phe Leu Gly
 35 40 45

Tyr Met Gln Tyr Ser Ile Val Asp Tyr Cys Trp Leu Ser Pro His Asn
 50 55 60

Ala Pro Val Leu Phe Leu Leu Trp Ser Gln Arg Xaa
 65 70 75

<210> 147

<211> 52

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (52)

<223> Xaa equals stop translation

<400> 147

Met Ala Gly Trp Phe Arg Gly Phe Phe Gly Phe Leu Phe Phe Phe Leu

```

1          5          90          15
Cys Leu Phe Asn Leu Lys Leu Phe Lys Leu Lys His Ser Gln Met Phe
      20          25          30
Gly Gly Lys His Pro Leu Lys Met Gly Pro Cys Ala Cys Leu Leu Gly
      35          40          45
Arg Arg Ser Xaa
      50
<210> 148
<211> 209
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (3)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (39)
<223> Xaa equals any of the naturally occurring L-amino acids
<400> 148
Met Ala Xaa Ser Ser Arg Gly Asn Ala Asp Ser Ile Val Ala Ser Leu
  1          5          10          15
Val Leu Met Val Leu Tyr Leu Ile Lys Lys Arg Leu Val Ala Cys Ala
      20          25          30
Ala Val Phe Tyr Gly Phe Xaa Val His Met Lys Ile Tyr Pro Val Thr
      35          40          45
Tyr Ile Leu Pro Ile Thr Leu His Leu Leu Pro Asp Arg Asp Asn Asp
      50          55          60
Lys Ser Leu Arg Gln Phe Arg Tyr Thr Phe Gln Ala Cys Leu Tyr Glu
      65          70          75          80
Leu Leu Lys Lys Leu Cys Asn Arg Ala Val Leu Leu Phe Val Ala Val
      85          90          95
Ala Gly Leu Thr Phe Phe Ala Leu Ser Phe Gly Phe Tyr Tyr Glu Tyr
      100          105          110
Gly Trp Glu Phe Leu Glu His Thr Tyr Phe Tyr His Leu Thr Arg Arg
      115          120          125
Asp Ile Arg His Asn Phe Ser Pro Tyr Phe Tyr Met Leu Tyr Leu Thr
      130          135          140
Ala Glu Ser Lys Trp Ser Phe Ser Leu Gly Ile Ala Ala Phe Leu Pro
      145          150          155          160
Gln Leu Ile Leu Leu Ser Ala Val Ser Phe Ala Tyr Tyr Arg Asp Leu
      165          170          175

```


Val Cys Thr Ser Gln Tyr Phe Leu Trp Val Pro Leu Ala Tyr Cys Leu
195 200 205

Lèu

```
<210> 149
<211> 219
<212> PRT
<213> Homo sapiens
```

```
<220> .
<221> SITE
<222> (168)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>
<221> SITE
<222> (174)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>  
<221> SITE  
<222> (198)  
<223> Xaa equals any of the naturally occurring L-amino acids
```

```

<220>
<221> SITE
<222> (213)
<223> Xaa equals any of the naturally occurring L-amino acids

```

```
<220>
<221> SITE
<222> (219)
<223> Xaa equals stop translation
```

<400> 149
Met Arg Ala Leu Leu Ala Leu Cys Leu Leu Leu Gly Trp Leu Arg Trp
1 5 10 15

Gly Pro Ala Gly Ala Gln Gln Ser Gly Glu Tyr Cys His Gly Trp Val
20 25 30

Asp Val Gln Gly Asn Tyr His Glu Gly Phe Gln Cys Pro Glu Asp Phe
35 40 45

Asp Thr Leu Asp Ala Thr Ile Cys Cys Gly Ser Cys Ala Leu Arg Tyr
50 55 60

Cys Cys Ala Ala Ala Asp Ala Arg Leu Glu Gln Gly Gly Cys Thr Asn
65 70 75 80

Asp Arg Arg Glu Leu Glu His Pro Gly Ile Thr Ala Gln Pro Val Tyr
85 90 95

92

Val	Pro	Phe	Leu	Ile	Val	Gly	Ser	Ile	Phe	Ile	Ala	Phe	Ile	Ile	Leu
			100					105					110		
Gly	Ser	Val	Val	Ala	Ile	Tyr	Cys	Cys	Thr	Cys	Leu	Arg	Pro	Lys	Glu
		115					120					125			
Pro	Ser	Gln	Gln	Pro	Ile	Arg	Phe	Ser	Leu	Arg	Ser	Tyr	Gln	Thr	Glu
	130					135					140				
Thr	Leu	Pro	Met	Ile	Leu	Thr	Ser	Thr	Ser	Pro	Arg	Ala	Pro	Ser	Arg
145					150					155					160
Gln	Ser	Ser	Thr	Ala	Thr	Ser	Xaa	Ser	Phe	Thr	Gly	Gly	Xaa	Ile	Arg
				165					170					175	
Arg	Phe	Phe	Ser	Ala	Ile	Trp	Phe	Pro	Gly	Val	Thr	Pro	Val	Phe	Arg
			180					185					190		
Leu	Pro	Pro	Ser	Ala	Xaa	Ala	Pro	Thr	Gly	Trp	Glu	Glu	Leu	Ser	Arg
	195						200					205			
Leu	Ser	Val	Pro	Xaa	Asp	Thr	Pro	Arg	Pro	Xaa					
	210					215									

<210> 150

<211> 50

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (41)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (50)

<223> Xaa equals stop translation

<400> 150

Met	Gly	Ala	His	Ser	Phe	Gly	Phe	Gln	Leu	Phe	Met	Ser	Val	Ser	Val
1				5				10					15		

Leu	Trp	Gly	Arg	Leu	Cys	Leu	Tyr	Gly	Arg	Phe	Ser	Val	Ile	Thr	Phe
		20						25					30		

Ala	Ser	Pro	Pro	Thr	Thr	Phe	Met	Xaa	Ile	Gln	Cys	Cys	Ser	His	Cys
	35						40					45			

Ser	Xaa														
	50														

<210> 151

<211> 41

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

93

<222> (41)

<223> Xaa equals stop translation

<400> 151

Met His Ile His Leu Asp Thr Ser Ser Leu Lys Thr Leu His Leu Gly
 1 5 10 15

Thr Leu Phe Phe Leu Phe Tyr Leu Ala Leu Thr Gln Asn Glu Glu Asn
 20 25 30

Ile Cys Asp Gly Lys Val Thr Leu Xaa
 35 40

<210> 152

<211> 108

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (108)

<223> Xaa equals stop translation

<400> 152

Met Pro Ile Ile Val Leu Ile Leu Val Ser Leu Leu Ser Gln Leu Met
 1 5 10 15

Val Ser Asn Pro Pro Tyr Ser Leu Tyr Pro Arg Ser Gly Thr Gly Gln
 20 25 30

Thr Ile Lys Met Gln Thr Glu Asn Leu Gly Val Val Tyr Tyr Val Asn
 35 40 45

Lys Asp Phe Lys Asn Glu Tyr Lys Gly Met Leu Leu Gln Lys Val Glu
 50 55 60

Lys Ser Val Glu Glu Asp Tyr Val Thr Asn Ile Arg Asn Asn Cys Trp
 65 70 75 80

Lys Glu Arg Gln Gln Lys Thr Asp Met Gln Tyr Ala Ala Lys Val Tyr
 85 90 95

Arg Asp Asp Arg Leu Arg Arg Arg Gln Met Pro Xaa
 100 105

<210> 153

<211> 157

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (157)

<223> Xaa equals stop translation

<400> 153

Met Gln Ala Ser Leu Trp Glu Pro Pro Arg Ser Gly Leu Pro Leu Trp
 1 5 10 15

94

Ala Glu Gly Leu Thr Phe Phe Tyr Cys Tyr Met Leu Leu Leu Val Leu
 20 25 30

Pro Cys Val Ala Leu Ser Glu Val Ser Met Gln Gly Glu His Ile Ala
 35 40 45

Pro Gln Lys Met Met Leu Tyr Pro Val Leu Ser Leu Ala Thr Val Asn
 50 55 60

Val Val Ala Val Leu Ala Arg Ala Ala Asn Met Ala Leu Phe Arg Asp
 65 70 75 80

Ser Arg Val Ser Ala Ile Phe Val Gly Lys Asn Val Val Ala Leu Ala
 85 90 95

Thr Lys Ala Cys Thr Phe Leu Glu Tyr Arg Arg Gln Val Arg Asp Phe
 100 105 110

Pro Pro Pro Ala Leu Ser Leu Glu Leu Gln Pro Pro Pro Pro Gln Arg
 115 120 125

Asn Ser Val Pro Pro Pro Pro Pro Leu His Gly Pro Pro Gly Arg Pro
 130 135 140

His Met Ser Ser Pro Thr Arg Asp Pro Leu Asp Thr Xaa
 145 150 155

<210> 154

<211> 151

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (151)

<223> Xaa equals stop translation

<400> 154

Met Gly Tyr Leu Phe Phe Leu Leu Phe Met Ile Cys Trp Met Ile Tyr
 1 5 10 15

Gly Cys Ile Ser Tyr Trp Gly Leu His Cys Glu Thr Thr Tyr Thr Lys
 20 25 30

Asp Gly Phe Trp Thr Tyr Ile Thr Gln Ile Ala Thr Cys Ser Pro Trp
 35 40 45

Met Phe Trp Met Phe Leu Asn Ser Val Phe His Phe Met Trp Val Ala
 50 55 60

Val Leu Leu Met Cys Gln Met Tyr Gln Ile Ser Cys Leu Gly Ile Thr
 65 70 75 80

Thr Asn Glu Arg Met Asn Ala Arg Arg Tyr Lys His Phe Lys Val Thr
 85 90 95

Thr Thr Ser Ile Glu Ser Pro Phe Asn His Gly Cys Val Arg Asn Ile
 100 105 110

95

Ile Asp Phe Phe Glu Phe Arg Cys Cys Gly Leu Phe Arg Pro Val Ile
115 120 125

Val Asp Trp Thr Arg Gln Tyr Thr Ile Glu Tyr Asp Gln Ile Ser Gly
130 135 140

Ser Gly Tyr Gln Leu Val Xaa
145 150

<210> 155

<211> 71

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (71)

<223> Xaa equals stop translation

<400> 155

Met Ala Leu Thr Leu Leu Ile Gln Ile Ile Phe Leu Ala Leu Gly
1 5 10 15

Lys Ile Ser Phe Ile Phe Val Cys Cys Lys Asp Gly Phe Ala Arg Ile
20 25 30

Ser His Asp Gln Asp Lys Leu Pro Ile Gln Lys Pro Thr Asp Thr Asn
35 40 45

Tyr Ile Met Arg Lys Lys Cys Ile Gln Leu Gly His Ile Ser Phe Glu
50 55 60

Leu Phe Gly Leu Lys Ala Xaa
65 70

<210> 156

<211> 490

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (134)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (389)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 156

Met Leu Ala Leu Thr Phe Met Phe Met Val Leu Glu Val Val Val Ser
1 5 10 15

Arg Val Thr Ser Ser Leu Ala Met Leu Ser Asp Ser Phe His Met Leu
20 25 30

Ser Asp Val Leu Ala Leu Val Val Ala Leu Val Ala Glu Arg Phe Ala
35 40 45

Arg Arg Thr His Ala Thr Gln Lys Asn Thr Phe Gly Trp Ile Arg Ala
 50 55 60
 Glu Val Met Gly Ala Leu Val Asn Ala Ile Phe Leu Thr Gly Leu Cys
 65 70 75 80
 Phe Ala Ile Leu Leu Glu Ala Ile Glu Arg Phe Ile Glu Pro His Glu
 85 90 95
 Met Gln Gln Pro Leu Val Val Leu Gly Val Gly Val Ala Gly Leu Leu
 100 105 110
 Val Asn Val Leu Gly Leu Cys Leu Phe His His His Ser Gly Phe Ser
 115 120 125
 Gln Asp Ser Gly His Xaa His Ser His Gly Gly His Gly His Gly His
 130 135 140
 Gly Leu Pro Lys Gly Pro Arg Val Lys Ser Thr Arg Pro Gly Ser Ser
 145 150 155 160
 Asp Ile Asn Val Ala Pro Gly Glu Gln Gly Pro Asp Gln Glu Glu Thr
 165 170 175
 Asn Thr Leu Val Ala Asn Thr Ser Asn Ser Asn Gly Leu Lys Leu Asp
 180 185 190
 Pro Ala Asp Pro Glu Asn Pro Arg Ser Gly Asp Thr Val Glu Val Gln
 195 200 205
 Val Asn Gly Asn Leu Val Arg Glu Pro Asp His Met Glu Leu Glu Glu
 210 215 220
 Asp Arg Ala Gly Gln Leu Asn Met Arg Gly Val Phe Leu His Val Leu
 225 230 235 240
 Gly Asp Ala Leu Gly Ser Val Ile Val Val Val Asn Ala Leu Val Phe
 245 250 255
 Tyr Phe Ser Trp Lys Gly Cys Ser Glu Gly Asp Phe Cys Val Asn Pro
 260 265 270
 Cys Phe Pro Asp Pro Cys Lys Pro Phe Val Glu Ile Ile Asn Ser Thr
 275 280 285
 His Ala Ser Val Tyr Glu Ala Gly Pro Cys Trp Val Leu Tyr Leu Asp
 290 295 300
 Pro Thr Leu Cys Val Val Met Val Cys Ile Leu Leu Tyr Thr Thr Tyr
 305 310 315 320
 Pro Leu Leu Lys Glu Ser Ala Leu Ile Leu Leu Gln Thr Val Pro Lys
 325 330 335
 Gln Ile Asp Ile Arg Asn Leu Ile Lys Glu Leu Arg Asn Val Glu Gly
 340 345 350
 Val Glu Glu Val His Glu Leu His Val Trp Gln Leu Ala Gly Ser Arg

97

355

360

365

Ile Ile Ala Thr Ala His Ile Lys Cys Glu Asp Pro Thr Ser Tyr Met
 370 375 380

Glu Val Ala Lys Xaa Ile Lys Asp Val Phe His Asn His Gly Ile His
 385 390 395 400

Ala Thr Thr Ile Gln Pro Glu Phe Ala Ser Val Gly Ser Lys Ser Ser
 405 410 415

Val Val Pro Cys Glu Leu Ala Cys Arg Thr Gln Cys Ala Leu Lys Gln
 420 425 430

Cys Cys Gly Thr Leu Pro Gln Ala Pro Ser Gly Lys Asp Ala Glu Lys
 435 440 445

Thr Pro Ala Val Ser Ile Ser Cys Leu Glu Leu Ser Asn Asn Leu Glu
 450 455 460

Lys Lys Pro Arg Arg Thr Lys Ala Glu Asn Ile Pro Ala Val Val Ile
 465 470 475 480

Glu Ile Lys Asn Met Pro Lys Gln Thr Thr
 485 490

<210> 157

<211> 31

<212> PRT

<213> Homo sapiens

<400> 157

Met Gln Pro Cys Val Ile Ser Trp Glu Gln Cys Ser Phe Val Ser Pro
 1 5 10 15

Arg Gly Pro His Val Tyr Ile Cys Phe His Asp Gln Arg Arg Phe
 20 25 30

<210> 158

<211> 115

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (96)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (100)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 158

Met Leu Gly Leu Leu Gly Ser Thr Ala Leu Val Gly Trp Ile Thr Gly
 1 5 10 15

Ala Ala Val Ala Val Leu Leu Leu Leu Leu Leu Ala Thr Cys Leu
 20 25 30

98

Phe His Gly Arg Gln Asp Cys Asp Val Glu Arg Asn Arg Thr Ala Ala
 35 40 45
 Gly Gly Asn Arg Val Arg Arg Ala Gln Pro Trp Pro Phe Arg Arg Arg
 50 55 60
 Gly His Leu Gly Ile Phe His His His Arg His Pro Gly His Val Ser
 65 70 75 80
 His Val Pro Asn Val Gly Leu His His His His His Pro Arg His Xaa
 85 90 95
 Pro His His Xaa His His His His His Pro His Arg His His Pro Arg
 100 105 110
 His Ala Arg
 115

<210> 159

<211> 380

<212> PRT

<213> Homo sapiens

<400> 159

Met Lys Arg Ala Ser Ala Gly Gly Ser Arg Leu Leu Ala Trp Val Leu
 1 5 10 15
 Trp Leu Gln Ala Trp Gln Val Ala Ala Pro Cys Pro Gly Ala Cys Val
 20 25 30
 Cys Tyr Asn Glu Pro Lys Val Thr Thr Ser Cys Pro Gln Gln Gly Leu
 35 40 45
 Gln Ala Val Pro Val Gly Ile Pro Ala Ala Ser Gln Arg Ile Phe Leu
 50 55 60
 His Gly Asn Arg Ile Ser His Val Pro Ala Ala Ser Phe Arg Ala Cys
 65 70 75 80
 Arg Asn Leu Thr Ile Leu Trp Leu His Ser Asn Val Leu Ala Arg Ile
 85 90 95
 Asp Ala Ala Ala Phe Thr Gly Leu Ala Leu Leu Glu Gln Leu Asp Leu
 100 105 110
 Ser Asp Asn Ala Gln Leu Arg Ser Val Asp Pro Ala Thr Phe His Gly
 115 120 125
 Leu Gly Arg Leu His Thr Val His Leu Asp Arg Cys Gly Leu Gln Glu
 130 135 140
 Leu Gly Pro Gly Leu Phe Arg Gly Leu Ala Ala Leu Gln Tyr Leu Tyr
 145 150 155 160
 Leu Gln Asp Asn Ala Leu Gln Ala Leu Pro Asp Asp Thr Phe Arg Asp
 165 170 175
 Leu Gly Asn Leu Thr His Leu Phe Leu His Gly Asn Arg Ile Ser Ser

99

180	185	190
Val Pro Glu Arg Ala Phe Arg Gly Leu His Ser Leu Asp Arg Leu Leu		
195	200	205
Leu His Gln Asn Arg Val Ala His Val His Pro His Ala Phe Arg Asp		
210	215	220
Leu Gly Arg Leu Met Thr Leu Tyr Leu Phe Ala Asn Asn Leu Ser Ala		
225	230	235 240
Leu Pro Thr Glu Ala Leu Ala Pro Leu Arg Ala Leu Gln Tyr Leu Arg		
245	250	255
Leu Asn Asp Asn Pro Trp Val Cys Asp Cys Arg Ala Arg Pro Leu Trp		
260	265	270
Ala Trp Leu Gln Lys Phe Arg Gly Ser Ser Ser Glu Val Pro Cys Ser		
275	280	285
Leu Pro Gln Arg Leu Ala Gly Arg Asp Leu Lys Arg Leu Ala Ala Asn		
290	295	300
Asp Leu Gln Gly Cys Ala Val Ala Thr Gly Pro Tyr His Pro Ile Trp		
305	310	315 320
Thr Gly Arg Ala Thr Asp Glu Glu Pro Leu Gly Leu Pro Lys Cys Cys		
325	330	335
Gln Pro Asp Ala Ala Asp Lys Ala Ser Val Leu Glu Pro Gly Arg Pro		
340	345	350
Ala Ser Ala Gly Asn Ala Leu Lys Gly Pro Arg Ala Gly Arg Gly Gln		
355	360	365
Ala Arg Arg Glu Thr Val Phe Gly Pro Arg Glu His		
370	375	380

<210> 160

<211> 92

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (92)

<223> Xaa equals stop translation

<400> 160

Met Arg Leu Cys Val Thr Gly Pro Pro Val Phe Phe Phe Phe Leu Asn
1 5 10 15

Phe Phe Phe Phe Leu Cys Val Gly Ala Cys Leu Gly Asp Leu Lys Ile
20 25 30

Ser Arg Leu Val Tyr Leu Cys Lys Ala Cys Leu Arg Leu Glu Tyr Leu
35 40 45

Gly Lys Glu Ser Asp Ser Met Leu Ser Glu Phe Leu Lys Gly Gln Lys

100
50 55 60
Lys Asn Trp Arg Leu Leu Lys Cys Arg Phe Glu Val Ile Phe Leu Lys
65 70 75 80

Tyr Tyr Phe Gly Phe Cys Asp Ile Val Lys Asn Xaa
85 90

<210> 161

<211> 45

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (45)

<223> Xaa equals stop translation

<400> 161

Met Lys Lys His Thr Lys Cys Gln Trp Leu Lys Met Thr Ile Leu Phe
1 5 10 15

Leu Thr Val Met Lys Ile Gly Tyr Gly Thr Ser Ala Ser Cys Tyr Arg
20 25 30

Pro Glu Val Leu Gly Leu Leu Met Pro His Pro Leu Xaa
35 40 45

<210> 162

<211> 46

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (46)

<223> Xaa equals stop translation

<400> 162

Met Ser Cys Gly Cys Cys Phe Ile His Ile Tyr Asn Leu Leu Leu Ser
1 5 10 15

Leu Cys Tyr Gly Leu Gly Val Glu Arg Val Lys Phe Phe Thr Phe Ser
20 25 30

Ile Leu Lys Lys Glu Thr Met Leu Leu Asn Tyr Leu Phe Xaa
35 40 45

<210> 163

<211> 128

<212> PRT

<213> Homo sapiens

<400> 163

Met Leu Ser Ser Pro Ile Leu Ala Ser Gly Pro Ala Trp Leu Ala Cys
1 5 10 15

Ser Phe Ser His Val Gln Trp Trp Val Cys Leu Ile Ala Gln Val Gln
20 25 30

Phe	Ser	Ala	Ala	Thr	Val	Ser	Pro	Gly	Arg	Ala	Gly	Thr	Gly	Ala	Ala
		35					40					45			
Pro	Ser	Val	Pro	Ala	Val	Trp	Ala	Ala	Glu	Ala	Arg	Gly	Pro	Ser	Val
		50				55					60				
Pro	Ser	Thr	Leu	Gln	Gly	Ser	Pro	Val	Leu	Gln	Arg	Asp	Leu	Ala	Asn
		65			70					75					80
Pro	Pro	Pro	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys
				85					90					95	
Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys
			100					105					110		
Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Gly	Gly	Pro
		115					120					125			

```
<220>
<221> SITE
<222> (58)
<223> Xaa equals stop translation
```

```

<400> 164
Met His Pro Trp Arg Leu Ser Met Cys Pro Ala Cys Val Leu Ala Ala
 1             5             10             15
Leu Pro Ala Leu Cys Ser Cys Leu Cys Ser Pro Asp Ala Arg Pro Pro
      20             25             30
His Gly Trp Met Ser Met Pro Phe Thr Pro His Pro Leu Val Ser Arg
      35             40             45
Ala Met Pro Thr Cys His Pro Cys Ser Xaa
      50             55

```

```
<220>
<221> SITE
<222> (98)
<223> Xaa equals stop translation
```

<400> 165
Met Tyr Arg Ala Ile Asp Ser Phe Pro Arg Trp Arg Ser Tyr Phe Tyr
1 5 10 15

102

Phe Ile Thr Leu Ile Phe Phe Leu Ala Trp Leu Val Lys Asn Val Phe
 20 25 30

Ile Ala Val Ile Ile Glu Thr Phe Ala Glu Ile Arg Val Gln Phe Gln
 35 40 45

Gln Met Trp Gly Ser Arg Ser Ser Thr Thr Ser Thr Ala Thr Thr Gln
 50 55 60

Met Phe His Glu Asp Ala Ala Gly Gly Trp Gln Leu Val Ala Val Gly
 65 70 75 80

Cys Gln Gln Ala Pro Gly Thr Arg Pro Ser Leu Pro Pro Gly Ala Val
 85 90 95

Gln Xaa

<210> 166
 <211> 60
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (60)
 <223> Xaa equals stop translation

<400> 166
 Met Thr Ser Phe Cys Glu Met Leu Lys Gly Ser Ala Ala Gly Cys Leu
 1 5 10 15

Val Leu Leu Ala Phe Ala Phe Tyr Leu Ala Cys Ser Phe Ser His Lys
 20 25 30

Thr Lys Ser His Ser His Tyr Ala Leu Phe Ile Leu Gln Asp Tyr Leu
 35 40 45

Leu Gly Asn Phe Tyr Tyr Ile Pro Leu Ser Pro Xaa
 50 55 60

<210> 167
 <211> 43
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (43)
 <223> Xaa equals stop translation

<400> 167
 Met Ser Val Ala His Met His Ala Cys Val Phe Leu Cys Ala Cys Val
 1 5 10 15

Phe Cys Leu Ala Glu Asn Ala Leu Glu Ser Val Ile Ile Leu Cys Tyr
 20 25 30

Ser Tyr Asn Lys Asp Glu Val Arg Glu His Xaa

103

35

40

<210> 168
<211> 54
<212> PRT
<213> Homo sapiens

<400> 168
Met Lys Thr His Leu Leu Met Phe Leu Leu Ser Cys Met Ala Arg Cys
1 5 10 15

Thr Gly Ile Val Pro Lys Arg Pro Gln Pro Ala Phe Pro Leu Arg Gly
20 25 30

Arg Arg Arg Lys Asn Ser Phe Leu Phe Leu Leu Ser Phe Ser Ile Glu
35 40 45

Phe Leu Leu Cys Val Trp
50

<210> 169
<211> 53
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (11)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 169
Met Cys Lys Ala Val Cys Lys His Arg Leu Xaa Leu Phe Ala Val Ser
1 5 10 15

Ser Phe Ser Leu Gly Leu Gly Trp Val Cys Val Leu Val Leu Met Leu
20 25 30

Trp Pro Val Arg Leu Ser Leu Ala Pro Arg Pro Val Gln Leu Gln Gln
35 40 45

Arg Arg Ser His Cys
50

<210> 170
<211> 54
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (54)
<223> Xaa equals stop translation

<400> 170
Met Phe Thr Ala Pro Leu Phe Phe Phe Phe Phe Glu Ile Ile Asn
1 5 10 15

Ser Met Arg Asn Leu Gly Leu Asn Ile Cys Leu Leu Cys Leu Ile
20 25 30

104

Glu His His Ser Arg Pro Ser Val Cys Leu Pro Phe Thr Pro Lys Ile
 35 40 45

Leu Thr Lys Lys Phe Xaa
 50

<210> 171

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals stop translation

<400> 171

Met Leu Cys Phe Leu Pro Ile Pro Leu Leu Ser Ile Leu Ser Pro Gln
 1 5 10 15

Thr Gln Ala Ser Arg Leu Leu Asp Glu Thr Val Arg Arg Lys His Phe
 20 25 30

Leu Thr Tyr Pro Phe Gly Ile Ser Ser Ile Ile Thr Gln Ala Leu Leu
 35 40 45

Xaa

<210> 172

<211> 224

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (183)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (214)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 172

Met Val Leu Val Ala Leu Ile Leu Leu His Ser Ala Leu Ala Gln Ser
 1 5 10 15

Arg Arg Asp Phe Ala Pro Pro Gly Gln Gln Lys Arg Glu Ala Pro Val
 20 25 30

Asp Val Leu Thr Gln Ile Gly Arg Ser Val Arg Gly Thr Leu Asp Ala
 35 40 45

Trp Ile Gly Pro Glu Thr Met His Leu Val Ser Glu Ser Ser Ser Gln
 50 55 60

Val Leu Trp Ala Ile Ser Ser Ala Ile Ser Val Ala Phe Phe Ala Leu

65 70 105 75 80
Ser Gly Ile Ala Ala Gln Leu Leu Asn Ala Leu Gly Leu Ala Gly Asp
85 90 95
Tyr Leu Ala Gln Gly Leu Lys Leu Ser Pro Gly Gln Val Gln Thr Phe
100 105 110
Leu Leu Trp Gly Ala Gly Ala Leu Val Val Tyr Trp Leu Leu Ser Leu
115 120 125
Leu Leu Gly Leu Val Leu Ala Leu Leu Gly Arg Ile Leu Trp Gly Leu
130 135 140
Lys Leu Val Ile Phe Leu Ala Gly Phe Val Ala Leu Met Arg Ser Val
145 150 155 160
Pro Asp Pro Ser Thr Arg Ala Leu Leu Leu Leu Ala Leu Leu Ile Leu
165 170 175
Tyr Ala Leu Leu Ser Arg Xaa Thr Gly Ser Arg Ala Ser Gly Ala Gln
180 185 190
Leu Glu Ala Lys Val Arg Gly Leu Glu Arg Gln Val Glu Glu Leu Arg
195 200 205
Trp Arg Gln Arg Gln Xaa Ala Lys Gly Ala Arg Ser Val Glu Glu Glu
210 215 220

<210> 173

<211> 201

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (10)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (11)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (27)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (50)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (60)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (84)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (178)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (180)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (190)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (201)

<223> Xaa equals stop translation

<400> 173

Met	Leu	Gln	Arg	Met	Leu	Ile	Asp	Val	Xaa	Xaa	Phe	Leu	Phe	Leu	Phe
1				5					10					15	

Ala	Val	Trp	Met	Val	Ala	Phe	Gly	Val	Ala	Xaa	Gln	Gly	Ile	Leu	Arg
		20					25						30		

Gln	Asn	Glu	Gln	Arg	Trp	Arg	Trp	Ile	Phe	Arg	Ser	Val	Ile	Tyr	Glu
	35					40						45			

Pro	Xaa	Leu	Ala	Met	Phe	Gly	Gln	Val	Pro	Ser	Xaa	Val	Asp	Gly	Thr
	50					55					60				

Thr	Tyr	Asp	Phe	Ala	His	Cys	Thr	Phe	Thr	Gly	Asn	Glu	Ser	Lys	Pro
65					70					75				80	

Leu	Cys	Val	Xaa	Leu	Asp	Glu	His	Asn	Leu	Pro	Arg	Phe	Pro	Glu	Trp
			85						90					95	

Ile	Thr	Ile	Pro	Leu	Val	Cys	Ile	Tyr	Met	Leu	Ser	Thr	Asn	Ile	Leu
			100					105					110		

Leu	Val	Asn	Leu	Leu	Val	Ala	Met	Phe	Gly	Tyr	Thr	Val	Gly	Thr	Val
	115					120						125			

Gln	Glu	Asn	Asn	Asp	Gln	Val	Trp	Lys	Phe	Gln	Arg	Tyr	Phe	Leu	Val
	130					135					140				

Gln	Glu	Tyr	Cys	Ser	Arg	Leu	Asn	Ile	Pro	Phe	Pro	Phe	Ile	Val	Phe
145					150					155				160	

107
Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe Lys Cys Cys Cys Lys
165 170 175
Glu Xaa Asn Xaa Glu Ser Ser Val Cys Cys Ser Lys Met Xaa Thr Met
180 185 190
Arg Leu Trp His Gly Arg Val Ser Xaa
195 200

<210> 174
<211> 93
<212> PRT
<213> Homo sapiens

<400> 174
Met Pro Arg Ala Thr Leu Trp Gly His Leu Ser Pro Ala Trp Val Leu
1 5 10 15
Val Pro Trp Thr Pro Arg Ala Cys Gly Gln Ala Ala Pro Gly Arg Gly
20 25 30
His Val Ala Ser Asp His Lys Ser Gly Leu Pro Trp Pro Lys His Cys
35 40 45
Ser Cys Leu His Pro Arg Ala Ser Gln Pro Cys Leu Phe Ser Leu Asn
50 55 60
Ser Asn Arg Thr Val Phe Thr Ala Ile Gln Arg Val Ala Leu Gly Trp
65 70 75 80
Thr Phe Trp Val Gln Ala Asn Leu Val Pro Arg Cys Thr
85 90

<210> 175
<211> 404
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (77)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (96)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (98)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE
<222> (108)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (122)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (124)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (126)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (175)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (192)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (210)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (236)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (239)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (309)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (335)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (389)
<223> Xaa equals any of the naturally occurring L-amino acids

109

<400> 175

Met His Pro Ile Pro Ser Ser Phe Met Ile Lys Ala Val Ser Ser Phe
 1 5 10 15
 Leu Thr Ala Glu Glu Ala Ser Val Gly Asn Pro Glu Gly Ala Phe Met
 20 25 30
 Lys Val Leu Gln Ala Arg Lys Asn Xaa Thr Ser Thr Glu Leu Ile Val
 35 40 45
 Glu Pro Glu Glu Pro Ser Asp Ser Ser Gly Ile Asn Leu Ser Gly Phe
 50 55 60
 Gly Ser Glu Gln Leu Asp Thr Asn Asp Glu Ser Asp Xaa Ile Ser Thr
 65 70 75 80
 Leu Ser Tyr Ile Leu Pro Tyr Phe Ser Ala Val Asn Leu Asp Val Xaa
 85 90 95
 Ser Xaa Leu Leu Pro Phe Ile Lys Leu Pro Thr Xaa Gly Asn Ser Leu
 100 105 110
 Ala Lys Ile Gln Thr Val Gly Gln Asn Xaa Gln Xaa Val Xaa Arg Val
 115 120 125
 Leu Met Gly Pro Arg Ser Ile Gln Lys Arg His Phe Lys Glu Val Gly
 130 135 140
 Arg Gln Ser Ile Arg Arg Glu Gln Gly Ala Gln Ala Ser Val Glu Asn
 145 150 155 160
 Ala Ala Glu Glu Lys Arg Leu Gly Ser Pro Ala Pro Arg Glu Xaa Glu
 165 170 175
 Gln Pro His Thr Gln Gln Gly Pro Glu Lys Leu Ala Gly Asn Ala Xaa
 180 185 190
 Tyr Thr Lys Pro Ser Phe Thr Gln Glu His Lys Ala Ala Val Ser Val
 195 200 205
 Leu Xaa Pro Phe Ser Lys Gly Ala Pro Ser Thr Ser Ser Pro Ala Lys
 210 215 220
 Ala Leu Pro Gln Val Arg Asp Arg Trp Lys Asp Xaa Thr His Xaa Ile
 225 230 235 240
 Ser Ile Leu Glu Ser Ala Lys Ala Arg Val Thr Asn Met Lys Ala Ser
 245 250 255
 Lys Pro Ile Ser His Ser Arg Lys Lys Tyr Arg Phe His Lys Thr Arg
 260 265 270
 Ser Arg Met Thr His Arg Thr Pro Lys Val Lys Lys Ser Pro Lys Phe
 275 280 285
 Arg Lys Lys Ser Tyr Leu Ser Arg Leu Met Leu Ala Asn Arg Pro Pro
 290 295 300
 Phe Ser Ala Ala Xaa Ser Leu Ile Asn Ser Pro Ser Gln Gly Ala Phe

Ile Asp Glu Tyr Met Arg Asn Phe Ser Asp Leu Arg Asn Gly Met Asp
115 120 125

111

Arg Ser Gly Ser Thr Ala Val Gly Val Met Ile Ser Pro Lys His Ile
 130 135 140
 Tyr Phe Ile Asn Cys Gly Asp Ser Arg Ala Val Leu Tyr Arg Asn Gly
 145 150 155 160
 Gln Val Cys Phe Ser Thr Gln Asp His Lys Pro Cys Asn Pro Arg Glu
 165 170 175
 Lys Glu Arg Ile Gln Asn Ala Gly Gly Ser Val Met Ile Gln Arg Val
 180 185 190
 Asn Gly Ser Leu Ala Val Ser Arg Ala Leu Gly Asp Tyr Asp Tyr Lys
 195 200 205
 Cys Val Asp Gly Lys Gly Pro Thr Glu Gln Leu Val Ser Pro Glu Pro
 210 215 220
 Glu Val Tyr Xaa Ile Leu Arg Ala Glu Glu Asp Glu Phe Ile Ile Leu
 225 230 235 240
 Ala Cys Asp Gly Ile Trp Asp Val Met Ser Asn Glu Glu Leu Cys Glu
 245 250 255
 Tyr Val Lys Ser Arg Leu Glu Val Ser Asp Asp Leu Glu Asn Val Cys
 260 265 270
 Asn Trp Val Val Asp Thr Cys Leu His Lys Gly Ser Arg Asp Asn Met
 275 280 285
 Ser Ile Val Leu Val Cys Phe Ser Asn Ala Pro Lys Val Ser Asp Glu
 290 295 300
 Ala Val Lys Lys Asp Ser Glu Leu Asp Lys His Leu Glu Ser Arg Val
 305 310 315 320
 Glu Glu Ile Met Glu Lys Ser Gly Glu Glu Gly Met Pro Asp Leu Ala
 325 330 335
 His Val Met Arg Ile Leu Ser Ala Glu Asn Ile Pro Asn Leu Pro Pro
 340 345 350
 Gly Gly Gly Leu Ala Gly Xaa Arg Asn Val Ile Glu Ala Val Tyr Ser
 355 360 365
 Arg Leu Asn Pro His Arg Glu Ser Asp Gly Gly Ala Gly Asp Leu Glu
 370 375 380
 Asp Pro Trp
 385
 <210> 177
 <211> 145
 <212> PRT
 <213> Homo sapiens
 <400> 177
 Met Ala Phe Phe Thr Gly Leu Trp Gly Pro Phe Thr Cys Val Ser Arg

112
 1 5 10 15
 Val Leu Ser His His Cys Phe Ser Thr Thr Gly Ser Leu Ser Ala Ile
 20 25 30
 Gln Lys Met Thr Arg Val Arg Val Val Asp Asn Ser Ala Leu Gly Asn
 35 40 45
 Ser Pro Tyr His Arg Ala Pro Arg Cys Ile His Val Tyr Lys Lys Asn
 50 55 60
 Gly Val Gly Lys Val Gly Asp Gln Ile Leu Leu Ala Ile Lys Gly Gln
 65 70 75 80
 Lys Lys Lys Ala Leu Ile Val Gly His Cys Met Pro Gly Pro Arg Met
 85 90 95
 Thr Pro Arg Phe Asp Ser Asn Asn Val Val Leu Ile Glu Asp Asn Gly
 100 105 110
 Asn Pro Val Gly Thr Arg Ile Lys Thr Pro Ile Pro Thr Ser Leu Arg
 115 120 125
 Lys Arg Glu Gly Glu Tyr Ser Lys Val Leu Ala Ile Ala Gln Asn Phe
 130 135 140

Val
145

<210> 178

<211> 140

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (129)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (132)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (134)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 178

Met Phe Phe Ser Leu Pro Gly Leu Trp Gln Ile Ala Ser Phe Thr His
 1 5 10 15

Asn Leu Ile Phe His Leu Trp Val Trp Gly Ser Glu Ser Gly Glu His
 20 25 30

Leu Gln Ser His Asn Asp Pro Asp Thr Arg Gln Gly Gly His Ile Pro
 35 40 45

113

Ile Arg Leu Leu Gly Glu Ser Ser Ala Ser Val Pro Gly Ser Ser Glu
50 55 60

Gly His Thr Gly Gly Pro Ala Pro Pro Arg Val Gly Gly Ser Ala Gly
65 70 75 80

Ile Ile Arg Thr His Val Val Phe Leu Val Ser Trp Pro Leu Leu Gln
85 90 95

Arg Glu Gln His Arg Leu Ser Trp Lys Leu Pro Ser Val Met Trp Gly
100 105 110

Asp Ser Arg Glu Pro His Leu Ala Arg Leu Asp Gln Ser Lys Trp Pro
115 120 125

Xaa Ala Thr Xaa Ala Xaa Gln Tyr Leu Gly Arg Gly
130 135 140

<210> 179
<211> 127
<212> PRT
<213> Homo sapiens

<400> 179

Met Val Pro Gly Ala Ala Gly Trp Cys Cys Leu Val Leu Trp Leu Pro
1 5 10 15

Ala Cys Val Ala Ala His Gly Phe Arg Ile His Asp Tyr Leu Tyr Phe
20 25 30

Gln Val Leu Ser Pro Gly Asp Ile Arg Tyr Ile Phe Thr Ala Thr Pro
35 40 45

Ala Lys Asp Phe Gly Gly Ile Phe His Thr Arg Tyr Glu Gln Ile His
50 55 60

Leu Val Pro Ala Glu Pro Pro Glu Ala Cys Gly Glu Leu Ser Asn Gly
65 70 75 80

Phe Phe Ile Gln Asp Gln Ile Ala Leu Val Glu Arg Gly Gly Cys Ser
85 90 95

Phe Leu Ser Lys Thr Arg Val Val Gln Glu His Gly Gly Arg Ala Val
100 105 110

Ile Ile Ser Asp Asn Ala Leu Thr Met Thr Ala Ser Thr Trp Arg
115 120 125

<210> 180
<211> 146
<212> PRT
<213> Homo sapiens

<400> 180

Met Gln Gln Ser Arg Leu Leu Leu Pro Phe Leu Phe Phe Leu Leu Glu
1 5 10 15

Gly Cys Ala Pro Ser Ser Leu Gly Pro Gly Ala Ala Pro Gly Ser Gly
20 25 30

114

His Ser Leu Gly Pro Pro Gly Ser Pro Gly Ala Pro Gly Pro Gln Pro
 35 40 45

Ala Val Gly Pro Ser Ser Pro Cys Gln Pro Gly Pro Ser Pro Ser Ser
 50 55 60

Pro Ala Ala Ala Ala Ala Ser Ser Gln Ser Ser Val Ala Ser Trp Pro
 65 70 75 80

Cys Thr Leu Arg Cys Ala Ala Pro Ser Pro Asp Ala Ser Ala Leu Arg
 85 90 95

Pro Ala Ala Ser Pro Ala Ala Thr Pro Ala Trp Ser Pro Gly Ser Gly
 100 105 110

Thr Ile Arg Val Leu Arg Pro Pro Ala Pro Ala Ala Ala Pro Ala Thr
 115 120 125

Ala Ile Thr Asn Arg Gly Pro Pro Arg Arg Arg Arg Arg Asn Ala Arg
 130 135 140

Thr Ala
 145

<210> 181

<211> 68

<212> PRT

<213> Homo sapiens

<400> 181

Met Lys Pro Thr Arg Ser Leu Trp Ile Ser Phe Leu Met Cys Cys Trp
 1 5 10 15

Ile Trp Phe Ala Asn Ile Leu Leu Arg Ile Phe Ala Ser Val Phe Phe
 20 25 30

Arg Asp Ile Gly Leu Lys Phe Ser Phe Phe Cys Cys Val Ser Ala Arg
 35 40 45

Leu Trp Tyr Gln Asp Asp Ala Gly Leu Ile Asn Glu Leu Gly Arg Ile
 50 55 60

Pro Ser Phe Tyr
 65

<210> 182

<211> 51

<212> PRT

<213> Homo sapiens

<400> 182

Met Thr Pro Val Phe Arg Ala Trp Gly Leu Trp Val Tyr Val Leu Pro
 1 5 10 15

Thr Gly Phe Pro Gly Pro Cys Cys Met Met Leu Leu Glu Leu Phe Pro
 20 25 30

Lys Glu Ser Val Pro Gln Ala Tyr Gln Gly Ile Leu Leu Tyr Leu His

115

35 40 45

Phe Gly Phe
50

<210> 183
<211> 85
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (68)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 183
Met Gly Met Pro Leu Val Thr Val Thr Ala Ala Thr Phe Pro Thr Leu
1 5 10 15

Ser Cys Pro Pro Arg Ala Trp Pro Glu Val Glu Ala Pro Glu Ala Pro
20 25 30

Ala Leu Pro Val Val Pro Glu Leu Pro Glu Val Pro Met Glu Met Pro
35 40 45

Leu Val Leu Pro Pro Glu Leu Glu Leu Leu Ser Leu Glu Ala Val His
50 55 60

Arg Tyr Gln Xaa Gly Gly Thr Leu Met Gly Trp Thr Arg Ala Glu Ala
65 70 75 80

Ser Ala Asn Gly Ser
85

<210> 184
<211> 191
<212> PRT
<213> Homo sapiens

<400> 184
Met Gly Asp His Leu Asp Leu Leu Leu Gly Val Val Leu Met Ala Gly
1 5 10 15

Pro Val Phe Gly Ile Pro Ser Cys Ser Phe Asp Gly Arg Ile Ala Phe
20 25 30

Tyr Arg Phe Cys Asn Leu Thr Gln Val Pro Gln Val Leu Asn Thr Thr
35 40 45

Glu Arg Leu Leu Leu Ser Phe Asn Tyr Ile Arg Thr Val Thr Ala Ser
50 55 60

Ser Phe Pro Phe Leu Glu Gln Leu Gln Leu Leu Glu Leu Gly Ser Gln
65 70 75 80

Tyr Thr Pro Leu Thr Ile Asp Lys Glu Ala Phe Arg Asn Leu Pro Asn
85 90 95

Leu Arg Ile Leu Asp Leu Gly Ser Ser Lys Ile Tyr Phe Leu His Pro

100 105 116 110
 Asp Ala Phe Gln Gly Leu Phe His Leu Phe Glu Leu Arg Leu Tyr Phe
 115 120 125
 Cys Gly Leu Ser Asp Ala Val Leu Lys Asp Gly Tyr Phe Arg Asn Leu
 130 135 140
 Lys Ala Leu Thr Arg Leu Asp Leu Ser Lys Asn Gln Ile Arg Ser Leu
 145 150 155 160
 Tyr Leu His Pro Ser Phe Gly Lys Leu Asn Ser Leu Lys Ser Ile Asp
 165 170 175
 Phe Ser Ser Asn Gln Ile Phe Leu Val Cys Glu His Glu Leu Glu
 180 185 190
 <210> 185
 <211> 231
 <212> PRT
 <213> Homo sapiens
 <400> 185
 Met Trp Ala Leu Gln Leu Ser Leu Pro Thr Cys Gly Leu Ala Ala Leu
 1 5 10 15
 Leu Thr His Met Arg Pro Cys Ser Ser Pro Tyr Pro His Ala Gly Leu
 20 25 30
 Ala Ala Leu Leu Thr His Met Gly Pro Cys Arg Ser Pro Tyr Pro His
 35 40 45
 Gly Gly Leu Ala Ala Val Leu Thr His Met Arg Ala Leu Gln Leu Ser
 50 55 60
 Leu Pro Thr Trp Gly Leu Ala Ala Leu Leu Thr His Met Arg Pro Cys
 65 70 75 80
 Ser Ser Pro Tyr Pro His Ala Gly Leu Ala Cys Cys Trp Leu Trp Ser
 85 90 95
 Leu Ser Ser His Arg Ser Leu Gln Val Gln Ala Thr His Arg Leu Val
 100 105 110
 Val Arg Thr Ile Lys Asp Arg Val Met Leu Lys Val Leu Pro Gln Thr
 115 120 125
 Arg Arg Arg Gly Pro Phe Leu Ser Ser Cys Arg Asn Asp Val Met Arg
 130 135 140
 Asn Cys Val Pro Arg His Ala Val Leu Val Thr Thr Cys Val Phe Val
 145 150 155 160
 Ser Phe Pro Thr His Cys Lys Val Gly Ile Thr Gly Pro Ile Thr Gln
 165 170 175
 Val Lys Gln Lys Pro Gly Asn His Ser Ser Pro Cys Pro Val Ile Gln
 180 185 190

117

Leu Val Ala Lys Ala Glu Phe Glu Leu Met Leu Pro Ser Val Pro Lys
 195 200 205

Pro Val Tyr Leu Thr Leu Val Leu Ser Cys Trp Cys Leu Cys Asp Val
 210 215 220

Pro Cys Leu Ser Val Ser Leu
 225 230

<210> 186

<211> 68

<212> PRT

<213> Homo sapiens

<400> 186

Met Tyr Leu Glu Val Ala Val Arg Pro Phe Leu Ile Ile Val Ala Phe
 1 5 10 15

Leu Gly Leu Ser Phe Leu Ala Leu Gln Met Pro Phe Trp Gln Gly Ser
 20 25 30

Ala Val Gly His Leu Arg Ala Gly Gly Ala Gly Val Ala His Leu Ser
 35 40 45

Gln Ala Gly Ile Ile Gln Ala Pro Val His Ser Gly Arg Glu Gly Gln
 50 55 60

Pro Pro Pro Gly
 65

<210> 187

<211> 211

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (100)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (103)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 187

Met Gly Glu Ala Ser Pro Pro Ala Pro Ala Arg Arg His Leu Leu Val
 1 5 10 15

Leu Leu Leu Leu Leu Ser Thr Leu Val Ile Pro Ser Ala Ala Ala Pro
 20 25 30

Ile His Asp Ala Asp Ala Gln Glu Ser Ser Leu Gly Leu Thr Gly Leu
 35 40 45

Gln Ser Leu Leu Gln Gly Phe Ser Arg Leu Phe Leu Lys Gly Asn Leu
 50 55 60

Leu Arg Gly Ile Asp Ser Leu Phe Ser Ala Pro Met Asp Phe Arg Gly

```
<210> 188
<211> 90
<212> PRT
<213> Homo sapiens
```

```
<210> 189
<211> 62
<212> PRT
<213> Homo sapiens
```

<400> 189

119

Met Glu Leu Met Ala Leu Phe Phe Arg Thr Thr Val Ala Ala Met
 1 5 10 15

Ala Ser Arg Gly Ala Leu Ala Leu Phe Leu Arg Lys Ile Leu Ser Glu
 20 25 30

Ala Lys Phe Lys Leu Ser Leu Thr Pro Gln Pro Pro Gln Pro Phe Tyr
 35 40 45

Ile Tyr Met Ala Tyr Tyr Ser Glu Asn Phe Phe Leu Lys Phe
 50 55 60

<210> 190

<211> 295

<212> PRT

<213> Homo sapiens

<400> 190

Met Leu Cys Cys Trp Phe Pro Trp Arg Ile Leu Ala Ala Gly Gln Val
 1 5 10 15

Pro Tyr Ser Pro His Ser Pro Gln Val Ala Gly Cys Asp Leu Thr Arg
 20 25 30

Cys Glu Ser Gly Gly Ala Arg Ala Leu Ser Ile Gln Arg Ala Ala Leu
 35 40 45

Val Val Leu Glu Asn Tyr Tyr Lys Asp Phe Thr Ile Tyr Asn Pro Asn
 50 55 60

Leu Leu Thr Ala Ser Lys Phe Arg Ala Ala Lys His Met Ala Gly Leu
 65 70 75 80

Lys Val Tyr Asn Val Asp Gly Pro Ser Asn Asn Ala Thr Gly Gln Ser
 85 90 95

Arg Ala Met Ile Ala Ala Ala Ala Arg Arg Arg Asp Ser Ser His Asn
 100 105 110

Glu Leu Tyr Tyr Glu Glu Ala Glu His Glu Arg Arg Val Lys Lys Arg
 115 120 125

Lys Ala Arg Leu Val Val Ala Val Glu Glu Ala Phe Ile His Ile Gln
 130 135 140

Arg Leu Gln Ala Glu Glu Gln Gln Lys Ala Pro Gly Glu Val Met Asp
 145 150 155 160

Pro Arg Glu Ala Ala Gln Ala Ile Phe Pro Ser Met Ala Arg Ala Leu
 165 170 175

Gln Lys Tyr Leu Arg Ile Thr Arg Gln Gln Asn Tyr His Ser Met Glu
 180 185 190

Ser Ile Leu Gln His Leu Ala Phe Cys Ile Thr Asn Gly Met Thr Pro
 195 200 205

Lys Ala Phe Leu Glu Arg Tyr Leu Ser Ala Gly Pro Thr Leu Gln Tyr
 210 215 220

120

Asp Lys Asp Arg Trp Leu Ser Thr Gln Trp Arg Leu Val Ser Asp Glu
 225 230 235 240

Ala Val Thr Asn Gly Leu Arg Asp Gly Ile Val Phe Val Leu Lys Cys
 245 250 255

Leu Asp Phe Ser Leu Val Val Asn Val Lys Lys Ile Pro Phe Ile Ile
 260 265 270

Leu Ser Glu Glu Phe Ile Asp Pro Lys Ser His Lys Phe Val Leu Arg
 275 280 285

Leu Gln Ser Glu Thr Ser Val
 290 295

<210> 191

<211> 295

<212> PRT

<213> Homo sapiens

<400> 191

Met Gly Leu Pro Val Ser Trp Ala Pro Pro Ala Leu Trp Val Leu Gly
 1 5 10 15

Cys Cys Ala Leu Leu Leu Ser Leu Trp Ala Leu Cys Thr Ala Cys Arg
 20 25 30

Arg Pro Glu Asp Ala Val Ala Pro Arg Lys Arg Ala Arg Arg Gln Arg
 35 40 45

Ala Arg Leu Gln Gly Ser Ala Thr Ala Ala Glu Ala Ser Leu Leu Arg
 50 55 60

Arg Thr His Leu Cys Ser Leu Ser Lys Ser Asp Thr Arg Leu His Glu
 65 70 75 80

Leu His Arg Gly Pro Arg Ser Ser Arg Ala Leu Arg Pro Ala Ser Met
 85 90 95

Asp Leu Leu Arg Pro His Trp Leu Glu Val Ser Arg Asp Ile Thr Gly
 100 105 110

Pro Gln Ala Ala Pro Ser Ala Phe Pro His Gln Glu Leu Pro Arg Ala
 115 120 125

Leu Pro Ala Ala Ala Ala Thr Ala Gly Cys Ala Gly Leu Glu Ala Thr
 130 135 140

Tyr Ser Asn Val Gly Leu Ala Ala Leu Pro Gly Val Ser Leu Ala Ala
 145 150 155 160

Ser Pro Val Val Ala Glu Tyr Ala Arg Val Gln Lys Arg Lys Gly Thr
 165 170 175

His Arg Ser Pro Gln Glu Pro Gln Gln Gly Lys Thr Glu Val Thr Pro
 180 185 190

Ala Ala Gln Val Asp Val Leu Tyr Ser Arg Val Cys Lys Pro Lys Arg

	195							121									205
Arg	Asp	Pro	Gly	Pro	Thr	Thr	Asp	Pro	Leu	Asp	Pro	Lys	Gly	Gln	Gly		
	210						215					220					
Ala	Ile	Leu	Ala	Leu	Ala	Gly	Asp	Leu	Ala	Tyr	Gln	Thr	Leu	Pro	Leu		
225					230					235					240		
Arg	Ala	Leu	Asp	Val	Asp	Ser	Gly	Pro	Leu	Glu	Asn	Val	Tyr	Glu	Ser		
				245					250					255			
Ile	Arg	Glu	Leu	Gly	Asp	Pro	Ala	Gly	Arg	Ser	Ser	Thr	Cys	Gly	Ala		
			260					265					270				
Gly	Thr	Pro	Pro	Ala	Ser	Ser	Cys	Pro	Ser	Leu	Gly	Arg	Gly	Trp	Arg		
		275					280					285					
Pro	Leu	Pro	Ala	Ser	Leu	Pro											
	290					295											
<210>	192																
<211>	338																
<212>	PRT																
<213>	Homo sapiens.																
<400>	192																
Met	Met	Arg	Thr	Cys	Val	Leu	Leu	Ser	Ala	Val	Leu	Trp	Cys	Leu	Thr		
1				5					10					15			
Gly	Val	Gln	Cys	Pro	Arg	Phe	Thr	Leu	Phe	Asn	Lys	Lys	Gly	Phe	Ile		
			20					25					30				
Tyr	Gly	Lys	Thr	Gly	Gln	Pro	Asp	Lys	Ile	Tyr	Val	Glu	Leu	His	Gln		
		35					40					45					
Asn	Ser	Pro	Val	Leu	Ile	Cys	Met	Asp	Phe	Lys	Leu	Ser	Lys	Lys	Glu		
	50					55					60						
Ile	Val	Asp	Pro	Thr	Tyr	Leu	Trp	Ile	Gly	Pro	Asn	Glu	Lys	Thr	Leu		
65					70					75					80		
Thr	Gly	Asn	Asn	Arg	Ile	Asn	Ile	Thr	Glu	Thr	Gly	Gln	Leu	Met	Val		
				85					90					95			
Lys	Asp	Phe	Leu	Glu	Pro	Leu	Ser	Gly	Leu	Tyr	Thr	Cys	Thr	Leu	Ser		
		100						105					110				
Tyr	Lys	Thr	Val	Lys	Ala	Glu	Thr	Gln	Glu	Glu	Lys	Thr	Val	Lys	Lys		
		115					120					125					
Arg	Tyr	Asp	Phe	Met	Val	Phe	Ala	Tyr	Arg	Glu	Pro	Asp	Tyr	Ser	Tyr		
	130					135					140						
Gln	Met	Ala	Val	Arg	Phe	Thr	Thr	Arg	Ser	Cys	Ile	Gly	Arg	Tyr	Asn		
145					150					155					160		
Asp	Val	Phe	Phe	Arg	Val	Leu	Lys	Lys	Ile	Leu	Asp	Ile	Leu	Ile	Ser		
				165					170						175		

122
 Asp Leu Ser Cys His Val Ile Glu Pro Ser Tyr Lys Cys His Ser Val
 180 185 190
 Glu Ile Pro Glu His Gly Leu Ile His Glu Leu Phe Ile Ala Phe Gln
 195 200 205
 Val Asn Pro Phe Ala Pro Gly Trp Lys Gly Ala Cys Asn Gly Ser Val
 210 215 220
 Asp Cys Glu Asp Thr Thr Asn His Asn Ile Leu Gln Ala Arg Asp Arg
 225 230 235 240
 Ile Glu Asp Phe Phe Arg Ser Gln Ala Tyr Ile Phe Tyr His Asn Phe
 245 250 255
 Asn Lys Thr Leu Pro Ala Met His Phe Val Asp His Ser Leu Gln Val
 260 265 270
 Val Arg Leu Asp Ser Cys Arg Pro Gly Phe Gly Lys Asn Glu Arg Leu
 275 280 285
 His Ser Asn Cys Ala Ser Cys Cys Val Val Cys Ser Pro Ala Thr Phe
 290 295 300
 Ser Pro Asp Val Asn Val Thr Cys Gln Thr Cys Val Ser Val Leu Thr
 305 310 315 320
 Tyr Gly Ala Lys Ser Cys Pro Gln Thr Ser Asn Lys Asn Gln Gln Tyr
 325 330 335

Glu Asp

<210> 193
 <211> 78
 <212> PRT
 <213> Homo sapiens

<400> 193
 Met Gln Gln Arg Gly Ala Ala Gly Ser Arg Gly Cys Ala Leu Phe Pro
 1 5 10 15
 Leu Leu Gly Val Leu Phe Phe Gln Val Ser Ala Pro Ala Gly Tyr Ala
 20 25 30
 Pro Leu Pro Ala Gly Gly Leu Gly Lys Met Val Ala Phe Pro Val Pro
 35 40 45
 Gly Arg Gly Val Ser Arg Lys Pro Pro His Ser Ser Gly Lys Glu Gly
 50 55 60
 Gly Arg Glu Arg Asp Val Gly Thr Met Ser Ser Pro Pro Arg
 65 70 75

<210> 194
 <211> 181
 <212> PRT
 <213> Homo sapiens

123

<400> 194

Met Met Leu Met Pro Tyr Gly Ala Leu Ile Ile Gly Phe Val Cys Gly
 1 5 10 15

Ile Ile Ser Thr Leu Gly Phe Val Tyr Leu Thr Pro Phe Leu Glu Ser
 20 25 30

Arg Leu His Ile Gln Asp Thr Cys Gly Ile Asn Asn Leu His Gly Ile
 35 40 45

Pro Gly Ile Ile Gly Gly Ile Val Gly Ala Val Thr Ala Ala Ser Ala
 50 55 60

Ser Leu Glu Val Tyr Gly Lys Glu Gly Leu Val His Ser Phe Asp Phe
 65 70 75 80

Gln Gly Phe Asn Gly Asp Trp Thr Ala Arg Thr Gln Gly Lys Phe Gln
 85 90 95

Ile Tyr Gly Leu Leu Val Thr Leu Ala Met Ala Leu Met Gly Gly Ile
 100 105 110

Ile Val Gly Leu Ile Leu Arg Leu Pro Phe Trp Gly Gln Pro Ser Asp
 115 120 125

Glu Asn Cys Phe Glu Asp Ala Val Tyr Trp Glu Met Pro Glu Gly Asn
 130 135 140

Ser Thr Val Tyr Ile Pro Glu Asp Pro Thr Phe Lys Pro Ser Gly Pro
 145 150 155 160

Ser Val Pro Ser Val Pro Met Val Ser Pro Leu Pro Met Ala Ser Ser
 165 170 175

Val Pro Leu Val Pro
 180

<210> 195

<211> 79

<212> PRT

<213> Homo sapiens

<400> 195

Met Leu Ser Leu Asp Phe Leu Asp Asp Val Arg Arg Met Asn Lys Arg
 1 5 10 15

Gln Val Ser Leu Ser Val Leu Phe Phe Ser Trp Leu Phe Leu Ser Leu
 20 25 30

Arg Gly Cys Cys Cys Gly Ala Arg Arg Thr Pro Gly Phe Trp Cys Glu
 35 40 45

Gly Leu Ser Trp Ser Asp Thr Arg Val Ile Arg Phe Leu Trp Arg Leu
 50 55 60

Trp Pro Glu Ala Ala Leu Ser Ala Ser Leu Phe Leu Thr Pro Asn
 65 70 75

<210> 196

124

<211> 69

<212> PRT

<213> Homo sapiens

<400> 196

Met Glu Pro Arg Ser Phe Leu Leu Pro Glu Leu Gly Gly Arg Val Ser
 1 5 10 15

His Ile Pro Leu Gly Leu Thr Leu Val Phe Ala Cys Phe Leu Met Val
 20 25 30

Arg Glu Thr Ala Gly Gly Phe Ser Phe Arg Ala Gly Asp Leu Glu Glu
 35 40 45

Ile Ser Arg Lys Arg Thr Asn Val Leu Gly Ser Leu Arg Gly Thr Glu
 50 55 60

Leu Ile Gly Tyr Ile
 65

<210> 197

<211> 271

<212> PRT

<213> Homo sapiens

<400> 197

Met Thr Gln Gly Lys Leu Ser Val Ala Asn Lys Ala Pro Gly Thr Glu
 1 5 10 15

Gly Gln Gln Gln Val His Gly Glu Lys Lys Glu Ala Pro Ala Val Pro
 20 25 30

Ser Ala Pro Pro Ser Tyr Glu Glu Ala Thr Ser Gly Glu Gly Met Lys
 35 40 45

Ala Gly Ala Phe Pro Pro Ala Pro Thr Ala Val Pro Leu His Pro Ser
 50 55 60

Trp Ala Tyr Val Asp Pro Ser Ser Ser Ser Ser Tyr Asp Asn Gly Phe
 65 70 75 80

Pro Thr Gly Asp His Glu Leu Phe Thr Thr Phe Ser Trp Asp Asp Gln
 85 90 95

Lys Val Arg Arg Val Phe Val Arg Lys Val Tyr Thr Ile Leu Leu Ile
 100 105 110

Gln Leu Leu Val Thr Leu Ala Val Val Ala Leu Phe Thr Phe Cys Asp
 115 120 125

Pro Val Lys Asp Tyr Val Gln Ala Asn Pro Gly Trp Tyr Trp Ala Ser
 130 135 140

Tyr Ala Val Phe Phe Ala Thr Tyr Leu Thr Leu Ala Cys Cys Ser Gly
 145 150 155 160

Pro Arg Arg His Phe Pro Trp Glu Pro Asp Ser Pro Asp Arg Leu Tyr
 165 170 175

125

Pro Val His Gly Leu Pro His Trp Asp Ala Val Gln Leu Leu Gln His
 180 185 190

His Leu Arg Ala Ala Val Pro Gly His His Gly Pro Cys Leu Pro Leu
 195 200 205

Ser His Arg Leu Gln Leu Pro Asp Gln Val Arg Leu His Leu Leu Pro
 210 215 220

Gly Arg Ala Leu Arg Ala Ser His Asp Ser Phe Leu Gln Arg Thr His
 225 230 235 240

Pro Gly His Pro Pro Thr Leu Pro Ile Cys Ala Leu Ala Pro Cys Ser
 245 250 255

Leu Cys Ser Thr Gly Ser Gly Cys Ile Tyr Ile Val Pro Gly Thr
 260 265 270

<210> 198

<211> 51

<212> PRT

<213> Homo sapiens

<400> 198

Met Lys Cys Thr Ala Val Phe Ala Pro Ser Ala Trp Pro Asn Thr Leu
 1 5 10 15

Ser Leu Leu Val Ser Leu His Thr Val Met Cys Ile Asn Trp His Leu
 20 25 30

Val Ser Ala Ser His Met His Ile Gly Arg Ile Val Ile Leu Glu Gly
 35 40 45

Asp Gly Met
 50

<210> 199

<211> 71

<212> PRT

<213> Homo sapiens

<400> 199

Met Pro Asn Thr Phe His Thr Tyr Arg Pro Ile Leu Leu Leu Leu Leu
 1 5 10 15

Leu Pro Ser Ser Ser His Gln Asn Met Ile Val Ser Leu Pro Gln Asn
 20 25 30

Met Tyr Phe Leu Ile Ala Val Ala Lys Arg Leu Cys Ala Glu Ser Leu
 35 40 45

Ala Ser Asp Pro Ala Pro Cys Asn Leu Ser Ala Leu Gln Ala Lys Pro
 50 55 60

Arg Pro Arg Leu Arg His Tyr
 65 70

<210> 200

<211> 60

126

<212> PRT

<213> Homo sapiens

<400> 200

Met Leu Tyr Trp Gly Asn Val Ala Leu Val Leu Pro Thr Pro Tyr Leu
 1 5 10 15

His Leu Ser Leu Thr Leu Leu Leu Ser Pro Glu Trp Leu Gly Glu Met
 20 25 30

Gly Arg Gly Leu Pro Trp Pro Gly His Leu Val Ala Ala Trp Leu Asp
 35 40 45

His Ile Ala Asn Glu Leu Gly Arg Gly Ala Ile Phe
 50 55 60

<210> 201

<211> 143

<212> PRT

<213> Homo sapiens

<400> 201

Met Lys Trp Glu Arg Gly Ser Pro Met Val Leu Leu Ala Leu Val Tyr
 1 5 10 15

Asp Val Cys Cys Ala Ser Arg Arg Gly Gly Gln Ser His Pro Thr Ser
 20 25 30

Gly Ser Asp Val Leu Pro Leu Pro Val Pro Ala Leu Ala Gln Pro Ala
 35 40 45

Gln Pro Ser Arg Leu Asp Ala Cys Ala Lys Ala Arg Gly Ser Gln Arg
 50 55 60

Ala Ala Gly Trp Pro Arg Ala Gly Ser Arg Leu Gly Pro Ala Val Gly
 65 70 75 80

Arg Ala Ala Ser Pro Ser Ser Leu Gln Thr His Gly Ser Ser Ser Gln
 85 90 95

Ser Ser Arg Gln Leu Pro Gly Pro Glu Met Ser Ser Ser Pro Pro Trp
 100 105 110

Gly Gln Ala Leu Pro Trp Pro Ser Ser Val Asn Pro Ser Phe Leu Cys
 115 120 125

Ala Val Ser Gly Leu Leu Thr Val Val Cys Val Cys Ala Arg Leu
 130 135 140

<210> 202

<211> 148

<212> PRT

<213> Homo sapiens

<400> 202

Met Gln Phe Ile Leu Thr Gly Ile Thr Leu Ser Gly Tyr Leu Phe Thr
 1 5 10 15

Phe Ser Ala Cys Ala Val Leu Ser Ala Ser Ile Thr Val Trp Gly Leu

127

20

25

30

Met Glu Cys Leu Ile His Arg His Gly Ser His Thr Thr Glu His Leu
 35 40 45

Thr Arg Thr Leu Thr Ser Gln Gln Ser Ser Arg Gly His Leu Ser Leu
 50 55 60

Ser His Ser Thr Thr Gln Ser Asn Gln Pro Glu Arg Thr Leu Ala Leu
 65 70 75 80

Leu Thr Gly Gly Thr Ala Asp Leu Ser Val Trp Arg Gln His Ser Pro
 85 90 95

Lys Met Gly Ala Ile Phe Gln Asp Ala Val Phe Ala Leu Asp Ser Gln
 100 105 110

Ala Tyr Leu Trp Gly Ile Val Ser Asn Arg Glu Asn Ile Trp Val Leu
 115 120 125

Glu Gln Trp Pro Pro Pro Lys Gly Phe His Ser Cys Gln Glu Thr Pro
 130 135 140

Gln Glu Ser His
 145

<210> 203

<211> 36

<212> PRT

<213> Homo sapiens

<400> 203

Met Trp Thr Cys Pro Gly Ile Ala Ala Leu Val Leu Met Ile Val Pro
 1 5 10 15

Gly Cys Ser Leu Cys Pro Ala Gln Val Val His His Val Gly Gln Arg
 20 25 30

Glu Ser Pro Ser
 35

<210> 204

<211> 406

<212> PRT

<213> Homo sapiens

<400> 204

Met Ser Gly Ala Pro Thr Ala Gly Ala Ala Leu Met Leu Cys Ala Ala
 1 5 10 15

Thr Ala Val Leu Leu Ser Ala Gln Gly Gly Pro Val Gln Ser Lys Ser
 20 25 30

Pro Arg Phe Ala Ser Trp Asp Glu Met Asn Val Leu Ala His Gly Leu
 35 40 45

Leu Gln Leu Gly Gln Gly Leu Arg Glu His Ala Glu Arg Thr Arg Ser
 50 55 60

128

Gln Leu Ser Ala Leu Glu Arg Arg Leu Ser Ala Cys Gly Ser Ala Cys
65 70 75 80

Gln Gly Thr Glu Gly Ser Thr Asp Leu Pro Leu Ala Pro Glu Ser Arg
85 90 95

Val Asp Pro Glu Val Leu His Ser Leu Gln Thr Gln Leu Lys Ala Gln
100 105 110

Asn Ser Arg Ile Gln Gln Leu Phe His Lys Val Ala Gln Gln Gln Arg
115 120 125

His Leu Glu Lys Gln His Leu Arg Ile Gln His Leu Gln Ser Gln Phe
130 135 140

Gly Leu Leu Asp His Lys His Leu Asp His Glu Val Ala Lys Pro Ala
145 150 155 160

Arg Arg Lys Arg Leu Pro Glu Met Ala Gln Pro Val Asp Pro Ala His
165 170 175

Asn Val Ser Arg Leu His Arg Leu Pro Arg Asp Cys Gln Glu Leu Phe
180 185 190

Gln Val Gly Glu Arg Gln Ser Gly Leu Phe Glu Ile Gln Pro Gln Gly
195 200 205

Ser Pro Pro Phe Leu Val Asn Cys Lys Met Thr Ser Asp Gly Gly Trp
210 215 220

Thr Val Ile Gln Arg Arg His Asp Gly Ser Val Asp Phe Asn Arg Pro
225 230 235 240

Trp Glu Ala Tyr Lys Ala Gly Phe Gly Asp Pro His Gly Glu Phe Trp
245 250 255

Leu Gly Leu Glu Lys Val His Ser Ile Thr Gly Asp Arg Asn Ser Arg
260 265 270

Leu Ala Val Gln Leu Arg Asp Trp Asp Gly Asn Ala Glu Leu Leu Gln
275 280 285

Phe Ser Val His Leu Gly Gly Glu Asp Thr Ala Tyr Ser Leu Gln Leu
290 295 300

Thr Ala Pro Val Ala Gly Gln Leu Gly Ala Thr Thr Val Pro Pro Ser
305 310 315 320

Gly Leu Ser Val Pro Phe Ser Thr Trp Asp Gln Asp His Asp Leu Arg
325 330 335

Arg Asp Lys Asn Cys Ala Lys Ser Leu Ser Gly Gly Trp Trp Phe Gly
340 345 350

Thr Cys Ser His Ser Asn Leu Asn Gly Gln Tyr Phe Arg Ser Ile Pro
355 360 365

Gln Gln Arg Gln Lys Leu Lys Lys Gly Ile Phe Trp Lys Thr Trp Arg
370 375 380

129

Gly Arg Tyr Tyr Pro Leu Gln Ala Thr Thr Met Leu Ile Gln Pro Met
 385 390 395 400

Ala Ala Glu Ala Ala Ser
 405

<210> 205
 <211> 91
 <212> PRT
 <213> Homo sapiens

<400> 205
 Met Glu Lys Thr Leu Phe Leu Tyr His Tyr Leu Pro Ala Leu Thr Phe
 1 5 10 15

Gln Ile Leu Leu Leu Pro Val Val Leu Gln His Ile Ser Asp His Leu
 20 25 30

Cys Arg Ser Gln Leu Gln Arg Ser Ile Phe Ser Ala Leu Val Val Ala
 35 40 45

Trp Tyr Ser Ser Ala Cys His Val Ser Asn Thr Leu Arg Pro Leu Thr
 50 55 60

Tyr Gly Asp Lys Ser Leu Ser Pro His Glu Leu Lys Ala Leu Arg Trp
 65 70 75 80

Lys Asp Ser Trp Asp Ile Leu Ile Arg Lys His
 85 90

<210> 206
 <211> 101
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (23)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (29)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 206
 Met Leu Leu Phe Gly Leu Cys Trp Gly Pro Tyr Val Ala Thr Leu Leu
 1 5 10 15

Leu Ser Val Leu Ala Tyr Xaa Gln Arg Pro Pro Leu Xaa Pro Gly Thr
 20 25 30

Leu Leu Ser Leu Leu Ser Leu Gly Ser Ala Ser Ala Ala Val Pro
 35 40 45

Val Ala Met Gly Leu Gly Asp Gln Arg Tyr Thr Ala Pro Trp Arg Ala
 50 55 60

130
 Ala Ala Gln Arg Cys Leu Gln Gly Leu Trp Gly Arg Ala Ser Arg Asp
 65 70 75 80
 Ser Pro Gly Pro Ser Ile Ala Tyr His Pro Ser Ser Gln Ser Ser Val
 85 90 95
 Asp Leu Asp Leu Asn
 100

<210> 207
 <211> 50
 <212> PRT
 <213> Homo sapiens

<400> 207
 Met Ser Ala Gly Lys Trp Leu Leu Leu Val Ile Phe Arg Asp Leu Gly
 1 5 10 15
 Cys Gly Val Ser Arg Thr Ser Pro His Leu Arg Ser Gly Glu Glu Gly
 20 25 30
 Arg Ile Trp Ser Leu Leu Thr Ala Cys Ser Cys Cys Cys Leu Phe Val
 35 40 45
 Ile Phe
 50

<210> 208
 <211> 161
 <212> PRT
 <213> Homo sapiens

<400> 208
 Met Thr Ser Ala Leu Arg Gly Val Ala Asp Asp Gln Gly Gln His Pro
 1 5 10 15
 Leu Leu Lys Met Leu Leu His Leu Leu Ala Phe Ser Ser Ala Ala Thr
 20 25 30
 Gly His Leu Gln Ala Ser Val Leu Thr Gln Cys Leu Lys Val Leu Val
 35 40 45
 Lys Leu Ala Glu Asn Thr Ser Cys Asp Phe Leu Pro Arg Phe Gln Cys
 50 55 60
 Val Phe Gln Val Leu Pro Lys Cys Leu Ser Pro Glu Thr Pro Leu Pro
 65 70 75 80
 Ser Val Leu Leu Ala Val Glu Leu Leu Ser Leu Leu Ala Asp His Asp
 85 90 95
 Gln Leu Ala Pro Gln Leu Cys Ser His Ser Glu Gly Cys Leu Leu Leu
 100 105 110
 Leu Leu Tyr Met Tyr Ile Thr Ser Arg Pro Asp Arg Val Ala Leu Glu
 115 120 125
 Thr Gln Trp Leu Gln Leu Glu Gln Glu Val Val Trp Leu Leu Ala Lys
 130 135 140

Leu Gly Val Gln Glu Pro Leu Ala Pro Ser His Trp Leu Gln Leu Pro
 145 150 155 160

Val

<210> 209

<211> 227

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (67)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (170)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 209

Met Leu Gly Leu Leu Leu Cys Thr Pro Arg Ala Trp Leu Thr Leu
 1 5 10 15

Ser Gly Pro Val Cys Phe Gln Gly Arg Gly Pro Ser Glu Val Pro Gln
 20 25 30

Arg Pro Pro Gln Leu Trp Val Val Ser Ile Ser Val Leu Gln Gly Gln
 35 40 45

His Arg Gly Arg Ala Gly Pro Arg Asp Glu Gln Glu Arg Gly Arg Asp
 50 55 60

Gln His Xaa Leu Pro Ala His Gly Arg Leu His Leu Ser Pro Arg Pro
 65 70 75 80

Glu Pro Gly Cys Arg Pro Ala Cys Ala Ala Pro Gly Gly Gln Pro Gly
 85 90 95

Val Val Ser Gly Leu Pro Ala Leu Gly Gln Pro Arg Glu Ala Ser Ala
 100 105 110

Pro Cys His Ile Ser Arg Leu Arg Thr Ala Ser Leu Ala Val Val Met
 115 120 125

Gly Ala Glu Lys Gly Gly Ala Glu Met Arg Pro Trp Pro Ala Val Gln
 130 135 140

Ala Pro Ala Pro Leu Pro Ser Val Gly Gly Thr Pro Ile Cys Ala Pro
 145 150 155 160

Gly Cys Gly Ser Lys Asp Thr Val Pro Xaa Leu Gln Pro Ser Val Pro
 165 170 175

Lys Gly Arg Ala Glu Ser Gly Phe Val Ser Ala Arg Phe Leu Cys Pro
 180 185 190

132

His Pro Pro Arg Ser Leu Leu Cys Leu Gly Pro Gly Pro Ser Leu Ser
 195 200 205

Gly Leu Pro Gly Pro Pro Ile Pro Ala Leu Leu Gln Gly Pro Leu Gly
 210 215 220

Leu Gly Cys
 225

<210> 210

<211> 351

<212> PRT

<213> Homo sapiens

<400> 210

Met Leu Thr Leu Arg Ser Leu Leu Phe Trp Ser Leu Val Tyr Cys Tyr
 1 5 10 15

Cys Gly Leu Cys Ala Ser Ile His Leu Leu Lys Leu Leu Trp Ser Leu
 20 25 30

Gly Lys Gly Pro Ala Gln Thr Phe Arg Arg Pro Ala Arg Glu His Pro
 35 40 45

Pro Ala Cys Leu Ser Asp Pro Ser Leu Gly Thr His Cys Tyr Val Arg
 50 55 60

Ile Lys Asp Ser Gly Leu Arg Phe His Tyr Val Ala Ala Gly Glu Arg
 65 70 75 80

Gly Lys Pro Leu Met Leu Leu Leu His Gly Phe Pro Glu Phe Trp Tyr
 85 90 95

Ser Trp Arg Tyr Gln Leu Arg Glu Phe Lys Ser Glu Tyr Arg Val Val
 100 105 110

Ala Leu Asp Leu Arg Gly Tyr Gly Glu Thr Asp Ala Pro Ile His Arg
 115 120 125

Gln Asn Tyr Lys Leu Asp Cys Leu Ile Thr Asp Ile Lys Asp Ile Leu
 130 135 140

Asp Ser Leu Gly Tyr Ser Lys Cys Val Leu Ile Gly His Asp Trp Gly
 145 150 155 160

Gly Met Ile Ala Trp Leu Ile Ala Ile Cys Tyr Pro Glu Met Val Met
 165 170 175

Lys Leu Ile Val Ile Asn Phe Pro His Pro Asn Val Phe Thr Glu Tyr
 180 185 190

Ile Leu Arg His Pro Ala Gln Leu Leu Lys Ser Ser Tyr Tyr Tyr Phe
 195 200 205

Phe Gln Ile Pro Trp Phe Pro Glu Phe Met Phe Ser Ile Asn Asp Phe
 210 215 220

Lys Val Leu Lys His Leu Phe Thr Ser His Ser Thr Gly Ile Gly Arg
 225 230 235 240

Lys	Gly	Cys	Gln	Leu	Thr	Thr	Glu	Asp	Leu	Glu	Ala	Tyr	Ile	Tyr	Val	245	250	255
Phe	Ser	Gln	Pro	Gly	Ala	Leu	Ser	Gly	Pro	Ile	Asn	His	Tyr	Arg	Asn	260	265	270
Ile	Phe	Ser	Cys	Leu	Pro	Leu	Lys	His	His	Met	Val	Thr	Thr	Pro	Thr	275	280	285
Leu	Leu	Leu	Trp	Gly	Glu	Asn	Asp	Ala	Phe	Met	Glu	Val	Glu	Met	Ala	290	295	300
Glu	Val	Thr	Lys	Ile	Tyr	Val	Lys	Asn	Tyr	Phe	Arg	Leu	Thr	Ile	Leu	305	310	315
Ser	Glu	Ala	Ser	His	Trp	Leu	Gln	Gln	Asp	Gln	Pro	Asp	Ile	Val	Asn	325	330	335
Lys	Leu	Ile	Trp	Thr	Phe	Leu	Lys	Glu	Glu	Thr	Arg	Lys	Lys	Asp		340	345	350

```
<220>  
<221> SITE  
<222> (59)  
<223> Xaa equals any of the naturally occurring L-amino acids
```

```

<220>
<221> SITE
<222> (61)
<223> Xaa equals any of the naturally occurring L-amino acids

```

```
<220>
<221> SITE
<222> (84)
<223> Xaa equals any of the naturally occurring L-amino acids
```

<400> 211
Met Gly His Leu Pro His Ile Leu Ser Leu Gly Leu Phe Leu Thr Leu
1 5 10 15

Leu Met Phe Cys Ile Thr Lys Ser Asp Gly Gln Asn Lys Ile Tyr Arg
20 25 30

Cys Phe Lys Lys Ala Ser Pro Gln Val Ile Val Thr His Thr Lys Met
35 40 45

Arg Ile Ala Ala Ile Ile Cys Ser Tyr Trp Xaa Gly Xaa Ala Asn Leu
50 55 60

Gly Thr Arg Ile Lys Leu Gln Leu Asn Ser Ala Val Tyr Lys Ile Phe
65 70 75 80

Val Ser Leu Xaa Arg Lys Arg Lys Arg Thr Leu Ser Trp

85

134
90

<210> 212
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 212
 Met Phe Gln Gln Gly Trp Ser Ser Pro Leu Leu Thr Pro Ala Phe Thr
 1 5 10 15
 Ile Leu Pro Met Ser Ser Leu Leu Thr Ser Leu His Pro Ala Pro Arg
 20 25 30
 Leu Pro Thr Leu Leu Ala Ala Ser Ser Pro Gln Leu Ala Pro Leu Thr
 35 40 45
 Cys Cys Phe Gln Tyr Pro Phe Leu Leu Ser Ala Ser Ser Leu Gly Asp
 50 55 60
 Ile His Pro Ser Ser Arg Asp Phe Ser Cys His Ile Asn Ser Asn Val
 65 70 75 80
 Ser Glu Leu Tyr Phe Leu Pro Pro Thr Ser Val Ser Leu Asn Val Arg
 85 90 95
 Ile Phe Tyr Phe Gln
 100

<210> 213
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 213
 Met Gly Trp Leu Gly Arg Thr Cys Leu Ala His Ser His Leu Asp Phe
 1 5 10 15
 Ile Ser Gly Ala Leu Leu Leu Thr Phe Ala Tyr Phe Leu Val Phe Gln
 20 25 30
 Val Cys Pro Val Ile Asn Lys Trp Leu Tyr Asn Leu Asp Gln His Val
 35 40 45
 Val Lys Glu Leu Ile Ser Lys Cys Trp Arg Trp Glu Gly Thr Gly Thr
 50 55 60
 Leu Gln Lys Lys Ala Gln Asn Pro Pro Ser Pro Phe Val Phe His Phe
 65 70 75 80
 Pro Leu Pro His Ser Gly Thr Ser Pro Arg Pro Lys Ile Ser Phe Leu
 85 90 95
 Leu Lys

<210> 214
 <211> 81
 <212> PRT

135

<213> Homo sapiens

<400> 214

Met Trp Gly Gly Ser Val Phe Leu Lys Pro Lys Leu Leu Gln Ala Gly
 1 5 10 15

Gly Phe Leu His Phe Leu Phe Val Leu Phe Leu Thr Ala Asp Ser Val
 20 25 30

His Leu Ser Val Gly Gly Glu Leu Leu Leu Arg Thr Gly Phe Lys Arg
 35 40 45

His Ile Pro Val Thr Phe Lys Asn Leu His Gly Gly Arg Ser Phe Ser
 50 55 60

Arg Ser Val Gly Trp Ser Thr Leu Gly Pro Thr Thr Leu Arg Arg Gly
 65 70 75 80

Arg

<210> 215

<211> 188

<212> PRT

<213> Homo sapiens

<400> 215

Met Phe His Gln Ile Trp Ala Ala Leu Leu Tyr Phe Tyr Gly Ile Ile
 1 5 10 15

Leu Asn Ser Ile Tyr Gln Cys Pro Glu His Ser Gln Leu Thr Thr Leu
 20 25 30

Gly Val Asp Gly Lys Glu Phe Pro Glu Val His Leu Gly Gln Trp Tyr
 35 40 45

Phe Ile Ala Gly Ala Ala Pro Thr Lys Glu Glu Leu Ala Thr Phe Asp
 50 55 60

Pro Val Asp Asn Ile Val Phe Asn Met Ala Ala Gly Ser Ala Pro Met
 65 70 75 80

Gln Leu His Leu Arg Ala Thr Ile Arg Met Lys Asp Gly Leu Cys Val
 85 90 95

Pro Arg Lys Trp Ile Tyr His Leu Thr Glu Gly Ser Thr Asp Leu Arg
 100 105 110

Thr Glu Gly Arg Pro Asp Met Lys Thr Glu Leu Phe Ser Ser Ser Cys
 115 120 125

Pro Gly Gly Ile Met Leu Asn Glu Thr Gly Gln Gly Tyr Gln Arg Phe
 130 135 140

Leu Leu Tyr Asn Arg Ser Pro His Pro Pro Glu Lys Cys Val Glu Glu
 145 150 155 160

Phe Lys Ser Leu Thr Ser Cys Leu Asp Ser Lys Ala Phe Leu Leu Thr
 165 170 175

136

Pro Arg Asn Gln Glu Ala Cys Glu Leu Ser Asn Asn
 180 185

<210> 216

<211> 44

<212> PRT

<213> Homo sapiens

<400> 216

Met Gln Arg Thr Phe Lys Tyr Leu His Phe Tyr Ile Ile Arg Phe Val
 1 5 10 15

Ser Thr Tyr Ala Phe Ile Val Phe Phe Pro Phe Ser Ser Ser His Val
 20 25 30

Asn Gly Pro Cys Glu Lys Asn Ile Pro Leu Gly Lys
 35 40

<210> 217

<211> 515

<212> PRT

<213> Homo sapiens

<400> 217

Met Gly Ser Ala Pro Trp Ala Pro Val Leu Leu Leu Ala Leu Gly Leu
 1 5 10 15

Arg Gly Leu Gln Ala Gly Gly Glu Trp Arg Arg Pro Pro Ala His Ser
 20 25 30

Pro Val Pro Ala Pro Pro Leu Arg Phe Ala Ser Pro His Ser Pro Gln
 35 40 45

Ala Pro Asp Pro Gly Phe Gln Glu Arg Phe Phe Gln Gln Arg Leu Asp
 50 55 60

His Phe Asn Phe Glu Arg Phe Gly Asn Lys Thr Phe Pro Gln Arg Phe
 65 70 75 80

Leu Val Ser Asp Arg Phe Trp Val Arg Gly Glu Gly Pro Ile Phe Phe
 85 90 95

Tyr Thr Gly Asn Glu Gly Asp Val Trp Ala Phe Ala Asn Asn Ser Gly
 100 105 110

Phe Val Ala Glu Leu Ala Ala Glu Arg Gly Ala Leu Leu Val Phe Ala
 115 120 125

Glu His Arg Tyr Tyr Gly Lys Ser Leu Pro Phe Gly Ala Gln Ser Thr
 130 135 140

Gln Arg Gly His Thr Glu Leu Leu Thr Val Glu Gln Ala Leu Ala Asp
 145 150 155 160

Phe Ala Glu Leu Leu Arg Ala Leu Arg Arg Asp Leu Gly Ala Gln Asp
 165 170 175

Ala Pro Ala Ile Ala Phe Gly Gly Ser Tyr Gly Gly Met Leu Ser Ala

[illegible]

138

Lys Ala Ala Arg Arg Glu Gln Gln Pro Ala Leu Arg Gly Gly Pro Arg
 500 505 510

Leu Ser Leu
 515

<210> 218

<211> 522

<212> PRT

<213> Homo sapiens

<400> 218

Met Ala Ala Ala Met Pro Leu Ala Leu Leu Val Leu Leu Leu Leu Gly
 1 5 10 15

Pro Gly Gly Trp Cys Leu Ala Glu Pro Pro Arg Asp Ser Leu Arg Glu
 20 25 30

Glu Leu Val Ile Thr Pro Leu Pro Ser Gly Asp Val Ala Ala Thr Phe
 35 40 45

Gln Phe Arg Thr Arg Trp Asp Ser Glu Leu Gln Arg Glu Gly Val Ser
 50 55 60

His Tyr Arg Leu Phe Pro Lys Ala Leu Gly Gln Leu Ile Ser Lys Tyr
 65 70 75 80

Ser Leu Arg Glu Leu His Leu Ser Phe Thr Gln Gly Phe Trp Arg Thr
 85 90 95

Arg Tyr Trp Gly Pro Pro Phe Leu Gln Ala Pro Ser Asp Thr Asp His
 100 105 110

Tyr Phe Leu Arg Tyr Ala Val Leu Pro Arg Glu Val Val Cys Thr Glu
 115 120 125

Asn Leu Thr Pro Trp Lys Lys Leu Leu Pro Cys Ser Ser Lys Ala Gly
 130 135 140

Leu Ser Val Leu Leu Lys Ala Asp Arg Leu Phe His Thr Ser Tyr His
 145 150 155 160

Ser Gln Ala Val His Ile Arg Pro Val Cys Arg Asn Ala Arg Cys Thr
 165 170 175

Ser Ile Ser Trp Glu Leu Arg Gln Thr Leu Ser Val Val Phe Asp Ala
 180 185 190

Phe Ile Thr Gly Gln Gly Lys Lys Asp Trp Ser Leu Phe Arg Met Phe
 195 200 205

Ser Arg Thr Leu Thr Glu Pro Cys Pro Leu Ala Ser Glu Ser Arg Val
 210 215 220

Tyr Val Asp Ile Thr Thr Tyr Asn Gln Asp Asn Glu Thr Leu Glu Val
 225 230 235 240

His Pro Pro Pro Thr Thr Thr Tyr Gln Asp Val Ile Leu Gly Thr Arg
 245 250 255

139

Lys Thr Tyr Ala Ile Tyr Asp Leu Leu Asp Thr Ala Met Ile Asn Asn
 260 265 270
 Ser Arg Asn Leu Asn Ile Gln Leu Lys Trp Lys Arg Pro Pro Glu Asn
 275 280 285
 Glu Ala Pro Pro Val Pro Phe Leu His Ala Gln Arg Tyr Val Ser Gly
 290 295 300
 Tyr Gly Leu Gln Lys Gly Glu Leu Ser Thr Leu Leu Tyr Asn Thr His
 305 310 315 320
 Pro Tyr Arg Ala Phe Pro Val Leu Leu Leu Asp Thr Val Pro Trp Tyr
 325 330 335
 Leu Arg Leu Tyr Val His Thr Leu Thr Ile Thr Ser Lys Gly Lys Glu
 340 345 350
 Asn Lys Pro Ser Tyr Ile His Tyr Gln Pro Ala Gln Asp Arg Leu Gln
 355 360 365
 Pro His Leu Leu Glu Met Leu Ile Gln Leu Pro Ala Asn Ser Val Thr
 370 375 380
 Lys Val Ser Ile Gln Phe Glu Arg Ala Leu Leu Lys Trp Thr Glu Tyr
 385 390 395 400
 Thr Pro Asp Pro Asn His Gly Phe Tyr Val Ser Pro Ser Val Leu Ser
 405 410 415
 Ala Leu Val Pro Ser Met Val Ala Ala Lys Pro Val Asp Trp Glu Glu
 420 425 430
 Ser Pro Leu Phe Asn Ser Leu Phe Pro Val Ser Asp Gly Ser Asn Tyr
 435 440 445
 Phe Val Arg Leu Tyr Thr Glu Pro Leu Leu Val Asn Leu Pro Thr Pro
 450 455 460
 Asp Phe Ser Met Pro Tyr Asn Val Ile Cys Leu Thr Cys Thr Val Val
 465 470 475 480
 Ala Val Cys Tyr Gly Ser Phe Tyr Asn Leu Leu Thr Arg Thr Phe His
 485 490 495
 Ile Glu Glu Pro Arg Thr Gly Gly Leu Ala Lys Arg Leu Ala Asn Leu
 500 505 510
 Ile Arg Arg Ala Arg Gly Val Pro Pro Leu
 515 520
 <210> 219
 <211> 52
 <212> PRT
 <213> Homo sapiens
 <400> 219
 Met Lys Ser His Ile Ser Trp Arg Leu Cys Ser Leu Leu Leu Ile Leu

140
 1 5 10 15
 Phe Ser Leu Ile Leu Ser Ala Cys Phe Ile Ser Ala Arg Trp Ser Ser
 20 25 30
 Asn Ser Asp Ile Phe Phe Ser Ala Trp Ser Ile Gln Leu Leu Ile Leu
 35 40 45
 Val Tyr Ala Ser
 50

<210> 220
 <211> 73
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (24)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 220
 Met Gly Phe Trp Cys Gly Cys Pro Phe Cys Leu Leu Val Phe Leu Leu
 1 5 10 15
 Thr Val Arg Thr Arg Ser Phe Xaa Ser Val Gly Val Cys Trp Arg Ser
 20 25 30
 Thr Pro Asp Pro Leu Cys Leu Gly Ile Ser Ser Arg Ser Cys Arg Thr
 35 40 45
 Ala Asp Ile Gly Glu Gln Gln Met Leu Leu Pro Asp Arg Ser Ser Gly
 50 55 60
 Ser Phe Val Ser Glu Tyr Pro Ala Met
 65 70

<210> 221
 <211> 54
 <212> PRT
 <213> Homo sapiens

<400> 221
 Met Tyr Arg Phe Phe Leu Cys Val Asp Leu Ser Phe Gln Leu Leu Trp
 1 5 10 15
 Val Ile Pro Arg Ser Thr Val Thr Gly Thr Tyr Gly Lys Asp Ile Phe
 20 25 30
 Ser Leu Ala Gly Asn His His Thr Val Phe Gln Ser Ser Cys Thr Ile
 35 40 45
 Leu His Thr His Gln His
 50

<210> 222
 <211> 72
 <212> PRT
 <213> Homo sapiens

141

<400> 222

Met Ala Thr Ile Leu Leu Lys Leu Pro Ile Leu Ser Ala Met Ile Lys
 1 5 10 15

Lys Pro Leu Arg Asn Tyr Leu Lys Thr Ser Glu Thr Thr Met Glu Lys
 20 25 30

Ile Ile Ile Gln Lys Leu Val Ala Asn Leu Lys Phe Leu Pro Leu Gly
 35 40 45

Thr Leu Gln Leu Ala Met Met Ile Ala Asn Leu Ile Lys Lys Leu Phe
 50 55 60

Phe Pro Leu Val Lys Ala Ala Lys
 65 70

<210> 223

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (26)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (51)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (68)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 223

Met Tyr Leu Ala Val Tyr Leu Leu Leu Phe Leu Cys Ile Cys Phe Tyr
 1 5 10 15

Phe Ile Ala Leu Phe Ser His Ala Leu Xaa Pro His Cys Phe Asn Tyr
 20 25 30

Pro Gly Phe Ser Phe Asn Leu Val His Trp Ser Ser Leu Ile Pro Pro
 35 40 45

Leu Pro Xaa Phe Phe Phe Phe Asn Ser Phe Ser Asn Cys Ser Leu Phe
 50 55 60

Phe Pro Tyr Xaa Leu
 65

<210> 224

<211> 57

<212> PRT

<213> Homo sapiens

<220>

142

<221> SITE

<222> (57)

<223> Xaa equals stop translation

<400> 224

Met Ala Lys Thr Asp Phe Ser Ile Ile Leu Leu Lys Leu His Cys Leu
1 5 10 15

Phe Phe Phe Ser Val Ile Ser Val His Cys Ala Gln Ser Phe Ile Ser
20 25 30

Val Thr Gln Thr Glu Pro Ser Pro Ala Val Cys Ile Phe Pro Ala Val
35 40 45

Gly Ser Gly Leu Gly Pro Cys Asp Xaa
50 55

<210> 225

<211> 77

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (77)

<223> Xaa equals stop translation

<400> 225

Met Ala Gly Pro Trp Thr Phe Thr Leu Leu Cys Gly Leu Leu Ala Ala
1 5 10 15

Thr Leu Ile Gln Ala Thr Leu Ser Pro Thr Ala Val Leu Ile Leu Gly
20 25 30

Pro Lys Val Ile Lys Glu Lys Leu Thr Gln Glu Leu Lys Asp His Asn
35 40 45

Ala Thr Ser Ile Leu Gln Gln Leu Pro Leu Leu Ser Ala Met Arg Glu
50 55 60

Lys Pro Ala Gly Ala Ser Leu Cys Trp Ala Ala Trp Xaa
65 70 75

<210> 226

<211> 45

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (45)

<223> Xaa equals stop translation

<400> 226

Met Asp Leu Tyr Phe Phe Leu Leu Ala Gly Ile Gln Ala Val Thr Ala
1 5 10 15

Leu Leu Phe Val Trp Ile Ala Gly Arg Tyr Glu Arg Ala Ser Gln Gly
20 25 30

Pro Ala Ser His Ser Arg Phe Ser Arg Asp Arg Gly Xaa
 35 40 45

<210> 227
 <211> 102
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (47)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (98)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (102)
 <223> Xaa equals stop translation

<400> 227
 Met Ser Trp Val Gln Ala Thr Leu Leu Ala Arg Gly Leu Cys Arg Ala
 1 5 10 15

Trp Gly Gly Thr Cys Gly Ala Ala Leu Thr Gly Thr Ser Ile Ser Gln
 20 25 30

Val Pro Arg Arg Leu Pro Arg Gly Leu His Cys Ser Ala Leu Xaa Ile
 35 40 45

Ala Leu Asn Ser Pro Trp Phe Pro Ala His Arg Asn Pro Gly Arg Gly
 50 55 60

Pro Pro Arg Leu Trp Cys Pro Leu Arg Thr Cys Leu Gly Arg Arg Leu
 65 70 75 80

Val Gly Asn Gly Thr Arg Arg Ala Ser Cys Arg Arg Cys Arg Asn Leu
 85 90 95

Arg Xaa Gln Arg Ala Xaa
 100

<210> 228
 <211> 132
 <212> PRT
 <213> Homo sapiens

<400> 228
 Met Thr Tyr Phe Ser Gly Leu Leu Val Ile Leu Ala Phe Ala Ala Trp
 1 5 10 15

Val Ala Leu Ala Glu Gly Leu Gly Val Ala Val Tyr Ala Ala Ala Val
 20 25 30

Leu Leu Gly Ala Gly Cys Ala Thr Ile Leu Val Thr Ser Leu Ala Met

[illegible]

```
<210> 229
<211> 66
<212> PRT
<213> Homo sapiens
```

```

<400> 229
Met Thr Tyr Phe Ser Gly Leu Leu Val Ile Leu Ala Phe Ala Ala Trp
  1             5             10             15
Val Ala Leu Ala Glu Gly Leu Gly Val Ala Val Tyr Ala Ala Ala Val
          20             25             30
Leu Leu Gly Ala Gly Cys Ala Thr Ile Leu Val Thr Ser Leu Ala Met
      35             40             45
Thr Ala Asp Leu Ile Gly Pro His Thr Asn Ser Gly Leu Ser Cys Thr
      50             55             60
Ala Pro
      65

```

```
<210> 230
<211> 73
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (73)
<223> Xaa equals stop translation
```

```

<400> 230
Met Pro Trp Lys Arg Ala Val Val Leu Leu Met Leu Trp Phe Ile Gly
 1             5             10             15
Gln Ala Met Trp Leu Ala Pro Ala Tyr Val Leu Glu Phe Gln Gly Lys
      20             25             30
Asn Thr Phe Leu Phe Ile Trp Leu Ala Gly Leu Phe Phe Leu Leu Ile

```

145

35 40 45

Asn Cys Ser Ile Leu Ile Gln Ile Ile Ser His Tyr Lys Glu Glu Pro
 50 55 60

Leu Thr Glu Arg Ile Lys Tyr Asp Xaa
 65 70

<210> 231
 <211> 293
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (134)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 231

Met Leu Ala Leu Thr Phe Met Phe Met Val Leu Glu Val Val Val Ser
 1 5 10 15

Arg Val Thr Ser Ser Leu Ala Met Leu Ser Asp Ser Phe His Met Leu
 20 25 30

Ser Asp Val Leu Ala Leu Val Val Ala Leu Val Ala Glu Arg Phe Ala
 35 40 45

Arg Arg Thr His Ala Thr Gln Lys Asn Thr Phe Gly Trp Ile Arg Ala
 50 55 60

Glu Val Met Gly Ala Leu Val Asn Ala Ile Phe Leu Thr Gly Leu Cys
 65 70 75 80

Phe Ala Ile Leu Leu Glu Ala Ile Glu Arg Phe Ile Glu Pro His Glu
 85 90 95

Met Gln Gln Pro Leu Val Val Leu Gly Val Gly Val Ala Gly Leu Leu
 100 105 110

Val Asn Val Leu Gly Leu Cys Leu Phe His His His Ser Gly Phe Ser
 115 120 125

Gln Asp Ser Gly His Xaa His Ser His Gly Gly His Gly His Gly His
 130 135 140

Gly Leu Pro Lys Gly Pro Arg Val Lys Ser Thr Arg Pro Gly Ser Ser
 145 150 155 160

Asp Ile Asn Val Ala Pro Gly Glu Gln Gly Pro Asp Gln Glu Glu Thr
 165 170 175

Asn Thr Leu Val Ala Asn Thr Ser Asn Ser Asn Gly Leu Lys Leu Asp
 180 185 190

Pro Ala Asp Pro Glu Asn Pro Arg Ser Gly Asp Thr Val Glu Val Gln
 195 200 205

Val Asn Gly Asn Leu Val Arg Glu Pro Asp His Met Glu Leu Glu Glu

146

210	215	220
Asp Arg Ala Gly Gln Leu Asn Met Arg Gly Val Phe Leu His Val Leu		
225	230	235 240
Gly Asp Ala Leu Gly Ser Val Ile Val Val Val Asn Ala Leu Val Phe		
	245	250 255
Tyr Phe Ser Trp Lys Gly Cys Ser Glu Gly Asp Phe Cys Val Asn Pro		
	260	265 270
Cys Phe Pro Asp Pro Cys Lys Ala Phe Val Glu Ile Leu Ile Val Leu		
	275	280 285

Met His Gln Phe Met
290

<210> 232
<211> 55
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (55)
<223> Xaa equals stop translation

<400> 232
Met Lys Thr His Leu Leu Met Phe Leu Leu Ser Cys Met Ala Arg Cys
1 5 10 15
Thr Gly Ile Val Pro Lys Arg Pro Gln Pro Ala Phe Pro Leu Arg Gly
20 25 30
Arg Arg Arg Lys Asn Ser Phe Leu Phe Leu Leu Ser Phe Ser Ile Glu
35 40 45
Phe Leu Leu Cys Val Trp Xaa
50 55

<210> 233
<211> 47
<212> PRT
<213> Homo sapiens

<400> 233
Met Lys Thr His Leu Leu Met Phe Leu Leu Ser Cys Met Ala Arg Cys
1 5 10 15
Thr Gly Ile Val Pro Lys Arg Pro Gln Pro Ala Phe Pro Leu Arg Gly
20 25 30
Lys Glu Lys Lys Lys Leu Leu Phe Ile Phe Thr Phe Phe Gln His
35 40 45

<210> 234
<211> 54
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (54)
<223> Xaa equals stop translation

<400> 234
Met Cys Lys Ala Val Cys Lys His Arg Leu Arg Leu Phe Ala Val Ser
1 5 10 15
Ser Phe Ser Leu Gly Leu Gly Trp Val Cys Val Leu Val Leu Met Leu
20 25 30
Trp Pro Val Arg Leu Ser Leu Ala Xaa Arg Pro Val Gln Leu Gln Gln
35 40 45
Arg Arg Ser His Cys Xaa
50

<210> 235
<211> 70
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (70)
<223> Xaa equals stop translation

<400> 235
Met Ser Arg Lys Ser Leu Ala Phe Pro Ile Ile Cys Ser Tyr Leu Cys
1 5 10 15
Phe Leu Thr Val Ala Thr Cys Ser Ile Ala Cys Thr Thr Val Phe Phe
20 25 30
Ala Asn Leu Arg His Thr Arg Tyr Ile Cys Ile Glu Leu Ser Ala Leu
35 40 45
Glu Thr Ser Gly Val Ile Ser Pro Gln Ile Asn Asn Val Pro Glu Val
50 55 60
His Gly Lys Tyr Ser Xaa
65 70

<210> 236
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (69)
<223> Xaa equals stop translation

148

<400> 236

Met Lys Pro Thr Arg Ser Leu Trp Ile Ser Phe Leu Met Cys Cys Trp
 1 5 10 15

Ile Trp Phe Ala Asn Ile Leu Leu Arg Ile Phe Ala Ser Val Phe Phe
 20 25 30

Arg Asp Ile Gly Leu Lys Phe Ser Phe Phe Cys Cys Val Ser Ala Arg
 35 40 45

Leu Trp Tyr Gln Asp Asp Ala Gly Leu Ile Asn Glu Leu Gly Arg Ile
 50 55 60

Pro Ser Phe Tyr Xaa
 65

<210> 237

<211> 67

<212> PRT

<213> Homo sapiens

<400> 237

Met Gly Glu Ala Ser Pro Pro Ala Pro Ala Arg Arg His Leu Leu Val
 1 5 10 15

Leu Leu Leu Leu Leu Ser Thr Leu Val Ile Pro Ser Ala Ala Ala Pro
 20 25 30

Ile His Asp Ala Asp Ala Gln Glu Ser Ser Leu Gly Leu Thr Gly Leu
 35 40 45

Gln Ser Leu Leu Gln Gly Phe Ser Arg Leu Phe Leu Lys Val Thr Cys
 50 55 60

Phe Gly Ala
 65

<210> 238

<211> 90

<212> PRT

<213> Homo sapiens

<400> 238

Met Leu Val Val Ser Thr Val Ile Ile Val Phe Trp Glu Phe Ile Asn
 1 5 10 15

Ser Thr Glu Gly Ser Phe Leu Trp Ile Tyr His Ser Lys Asn Pro Glu
 20 25 30

Val Asp Asp Ser Ser Ala Gln Lys Gly Trp Trp Phe Leu Ser Trp Phe
 35 40 45

Asn Asn Gly Ile His Asn Tyr Gln Gln Gly Glu Glu Asp Ile Asp Lys
 50 55 60

Glu Lys Gly Arg Glu Glu Thr Lys Gly Arg Lys Met Thr Gln Gln Ser
 65 70 75 80

149

Phe Gly Tyr Gly Thr Gly Leu Ile Gln Thr
 85 90

<210> 239

<211> 140

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (117)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 239

Met Ala Phe Lys Leu Leu Ile Leu Leu Ile Gly Thr Trp Ala Leu Phe
 1 5 10 15

Phe Arg Lys Arg Arg Ala Asp Met Pro Arg Val Phe Val Phe Arg Ala
 20 25 30

Leu Leu Leu Val Leu Ile Phe Leu Phe Cys Gly Phe Pro Ile Gly Phe
 35 40 45

Phe Thr Gly Ser Ala Phe Trp Thr Leu Gly Asn Arg Asn Tyr Gln Gly
 50 55 60

Ile Val Gln Tyr Ala Val Ser Pro Cys Gly Met Pro Ser Ser Phe His
 65 70 75 80

Pro Leu Leu Ala Ile Arg Pro Cys Trp Ser Ser Gly Ser Leu Gln Pro
 85 90 95

Asn Val Pro Arg Cys Arg Leu Val Pro Leu Pro Thr Glu Trp Gly Asn
 100 105 110

Pro Arg Phe Gln Xaa Gly Thr Pro Glu Tyr Pro Ala Ser Ser Ile Gly
 115 120 125

Gly Pro Arg Lys Leu Leu Gln Arg Phe His His Leu
 130 135 140

<210> 240

<211> 37

<212> PRT

<213> Homo sapiens

<400> 240

Met Gly Leu Pro Val Ser Trp Ala Pro Pro Ala Leu Trp Val Leu Gly
 1 5 10 15

Cys Cys Ala Leu Leu Leu Ser Leu Trp Ala Leu Cys Thr Ala Cys Arg
 20 25 30

Ser Pro Arg Thr Leu
 35

<210> 241

<211> 21

<212> PRT

150

<213> Homo sapiens

<220>

<221> SITE

<222> (21)

<223> Xaa equals stop translation

<400> 241

Arg	Leu	Leu	Asn	Leu	Ser	Val	Pro	Met	Phe	Thr	Phe	Ile	Val	Val	Lys
1				5					10					15	

Arg	Tyr	Ala	Thr	Xaa
				20

<210> 242

<211> 138

<212> PRT

<213> Homo sapiens

<400> 242

Met	Ala	Tyr	Leu	Thr	Gly	Met	Leu	Ser	Ser	Tyr	Tyr	Asn	Thr	Thr	Ser
1				5					10					15	

Val	Leu	Leu	Cys	Leu	Gly	Ile	Thr	Ala	Leu	Val	Cys	Leu	Ser	Val	Thr
			20					25					30		

Val	Phe	Ser	Phe	Gln	Thr	Lys	Phe	Asp	Phe	Thr	Ser	Cys	Gln	Gly	Val
			35				40					45			

Leu	Phe	Val	Leu	Leu	Met	Thr	Leu	Phe	Phe	Ser	Gly	Leu	Ile	Leu	Ala
	50					55					60				

Ile	Leu	Leu	Pro	Phe	Gln	Tyr	Val	Pro	Trp	Leu	His	Ala	Val	Tyr	Ala
65					70					75				80	

Ala	Leu	Gly	Ala	Gly	Val	Phe	Thr	Leu	Phe	Leu	Ala	Leu	Asp	Thr	Gln
				85					90					95	

Leu	Leu	Met	Gly	Asn	Arg	Arg	His	Ser	Leu	Ser	Pro	Glu	Glu	Tyr	Ile
			100					105					110		

Phe	Gly	Ala	Leu	Asn	Ile	Tyr	Leu	Asp	Ile	Ile	Tyr	Ile	Phe	Thr	Phe
		115					120					125			

Phe	Leu	Gln	Leu	Phe	Gly	Thr	Asn	Arg	Glu
	130					135			

<210> 243

<211> 175

<212> PRT

<213> Homo sapiens

<400> 243

Met	Ala	Gln	Trp	Thr	Ser	Thr	Gly	Pro	Gly	Lys	Pro	Thr	Arg	Arg	Gly
1				5					10					15	

Leu	Gly	Ile	Pro	Thr	Ala	Ser	Ser	Gly	Trp	Val	Trp	Arg	Arg	Cys	Ile
			20					25						30	

151

Ala Ser Trp Gly Thr Ala Thr Ala Ala Trp Pro Cys Ser Cys Gly Thr
 35 40 45

Gly Met Ala Thr Pro Ser Cys Cys Ser Ser Pro Cys Thr Trp Val Ala
 50 55 60

Arg Thr Arg Pro Ile Ala Cys Ser Ser Leu His Pro Trp Pro Ala Ser
 65 70 75 80

Trp Ala Pro Pro Pro Ser His Pro Ala Ala Ser Pro Tyr Pro Ser Pro
 85 90 95

Leu Gly Thr Arg Ile Thr Thr Ser Ala Gly Thr Arg Thr Ala Pro Arg
 100 105 110

Ala Ser Leu Glu Ala Gly Gly Leu Ala Pro Ala Ala Ile Pro Thr Phe
 115 120 125

Asn Gly Pro Val Leu Pro Ala Pro Ser His Ser Ser Gly Arg Ser Leu
 130 135 140

Arg Arg Glu Ser Ser Gly Arg Pro Ala Gly Arg Tyr Tyr Pro Leu Gln
 145 150 155 160

Ala Thr Thr Met Leu Ile Gln Pro Met Ala Ala Glu Ala Ala Ser
 165 170 175

<210> 244

<211> 39

<212> PRT

<213> Homo sapiens

<400> 244

Met Leu Gly Leu Leu Leu Leu Cys Thr Pro Arg Ala Trp Leu Thr Leu
 1 5 10 15

Ser Gly Pro Val Cys Phe Gln Gly Arg Asp Pro Leu Arg Ser His Arg
 20 25 30

Gly His Pro Ser Cys Gly Ser
 35

<210> 245

<211> 47

<212> PRT

<213> Homo sapiens

<400> 245

Met Leu Ser Ile Ile Pro Asn Asp Arg Leu Phe Ile Asn Leu Ile Phe
 1 5 10 15

Leu Ser Asn Phe Leu Pro Ser Val Leu Trp Glu Pro Ala Gly Gln Met
 20 25 30

Trp Tyr Thr His Val Arg Tyr Pro Ser Gly Arg Leu Leu Ser Leu
 35 40 45

<210> 246

<211> 34

152

<212> PRT

<213> Homo sapiens

<400> 246

Met Thr Gly Phe Ala Gln Phe Cys Val Ile Leu Gly Leu Asn Leu Ser
 1 5 10 15

Leu Phe Gly Thr Phe Pro Tyr Leu Leu Pro Ser Ser Glu Ser Arg Cys
 20 25 30

Arg Lys

<210> 247

<211> 490

<212> PRT

<213> Homo sapiens

<400> 247

Met Gly Ser Ala Pro Trp Ala Pro Val Leu Leu Leu Ala Leu Gly Leu
 1 5 10 15

Arg Gly Leu Gln Ala Gly Ala Arg Ser Gly Pro Arg Leu Pro Gly Ala
 20 25 30

Leu Leu Pro Ala Ala Ser Gly Pro Leu Gln Leu Arg Ala Leu Arg Gln
 35 40 45

Gln Asp Leu Pro Ser Ala Leu Pro Gly Val Gly Gln Val Leu Gly Pro
 50 55 60

Gly Arg Gly Ala His Leu Leu Leu His Trp Glu Arg Gly Arg Arg Val
 65 70 75 80

Gly Leu Arg Gln Gln Leu Gly Leu Arg Arg Gly Leu Ala Ala Glu Arg
 85 90 95

Gly Ala Leu Leu Val Phe Ala Glu His Arg Tyr Tyr Gly Lys Ser Leu
 100 105 110

Pro Phe Gly Ala Gln Ser Thr Gln Arg Gly His Thr Glu Leu Leu Thr
 115 120 125

Val Glu Gln Ala Leu Ala Asp Phe Ala Glu Leu Leu Arg Ala Leu Arg
 130 135 140

Arg Asp Leu Gly Ala Gln Asp Ala Pro Ala Ile Ala Phe Gly Gly Ser
 145 150 155 160

Tyr Gly Gly Met Leu Ser Ala Tyr Leu Arg Met Lys Tyr Pro His Leu
 165 170 175

Val Ala Gly Ala Leu Ala Ala Ser Ala Pro Val Leu Ser Val Ala Gly
 180 185 190

Leu Gly Asp Ser Asn Gln Phe Phe Arg Asp Val Thr Ala Asp Phe Glu
 195 200 205

Gly Gln Ser Pro Lys Cys Thr Gln Gly Val Arg Glu Ala Phe Arg Gln

153

210	215	220
Ile Lys Asp Leu Phe Leu Gln Gly Ala Tyr Asp Thr Val Arg Trp Glu		
225	230	235 240
Phe Gly Thr Cys Gln Pro Leu Ser Asp Glu Lys Asp Leu Thr Gln Leu		
	245	250 255
Phe Met Phe Ala Arg Asn Ala Phe Thr Val Leu Ala Met Met Asp Tyr		
	260	265 270
Pro Tyr Pro Thr Asp Phe Leu Gly Pro Leu Pro Ala Asn Pro Val Lys		
	275	280 285
Val Gly Cys Asp Arg Leu Leu Ser Glu Ala Gln Arg Ile Thr Gly Leu		
	290	295 300
Arg Ala Leu Ala Gly Leu Val Tyr Asn Ala Ser Gly Ser Glu His Cys		
305	310	315 320
Tyr Asp Ile Tyr Arg Leu Tyr His Ser Cys Ala Asp Pro Thr Gly Cys		
	325	330 335
Gly Thr Gly Pro Asp Ala Arg Ala Trp Asp Tyr Gln Ala Cys Thr Glu		
	340	345 350
Ile Asn Leu Thr Phe Ala Ser Asn Asn Val Thr Asp Met Phe Pro Asp		
	355	360 365
Leu Pro Phe Thr Asp Glu Leu Arg Gln Arg Tyr Cys Leu Asp Thr Trp		
	370	375 380
Gly Val Trp Pro Arg Pro Asp Trp Leu Leu Thr Ser Phe Trp Gly Gly		
385	390	395 400
Asp Leu Arg Ala Ala Ser Asn Ile Ile Phe Ser Asn Gly Asn Leu Asp		
	405	410 415
Pro Trp Ala Gly Gly Gly Ile Arg Arg Asn Leu Ser Ala Ser Val Ile		
	420	425 430
Ala Val Thr Ile Gln Gly Gly Ala His His Leu Asp Leu Arg Ala Ser		
	435	440 445
His Pro Glu Asp Pro Ala Ser Val Val Glu Ala Arg Lys Leu Glu Ala		
	450	455 460
Thr Ile Ile Gly Glu Trp Val Lys Ala Ala Arg Arg Glu Gln Gln Pro		
465	470	475 480
Ala Leu Arg Gly Gly Pro Arg Leu Ser Leu		
	485	490

<210> 248

<211> 555

<212> PRT

<213> Homo sapiens

<220>

154

<221> SITE

<222> (555)

<223> Xaa equals stop translation

<400> 248

Gly Gly Gly Tyr Ala Leu Ala Leu Leu Val Leu Leu Leu Leu Gly Pro
 1 5 10 15

Gly Gly Trp Cys Leu Ala Glu Pro Pro Arg Asp Ser Leu Arg Glu Glu
 20 25 30

Leu Val Ile Thr Pro Leu Pro Ser Gly Asp Val Ala Ala Thr Phe Gln
 35 40 45

Phe Arg Thr Arg Trp Asp Ser Glu Leu Gln Arg Glu Gly Val Ser His
 50 55 60

Tyr Arg Leu Phe Pro Lys Ala Leu Gly Gln Leu Ile Ser Lys Tyr Ser
 65 70 75 80

Leu Arg Glu Leu His Leu Ser Phe Thr Gln Gly Phe Trp Arg Thr Arg
 85 90 95

Tyr Trp Gly Pro Pro Phe Leu Gln Ala Pro Ser Asp Thr Asp His Tyr
 100 105 110

Phe Leu Arg Tyr Ala Val Leu Pro Arg Glu Val Val Cys Thr Glu Asn
 115 120 125

Leu Thr Pro Trp Lys Lys Leu Leu Pro Cys Ser Ser Lys Ala Gly Leu
 130 135 140

Ser Val Leu Leu Lys Ala Asp Arg Leu Phe His Thr Ser Tyr His Ser
 145 150 155 160

Gln Ala Val His Ile Arg Pro Val Cys Arg Asn Ala Arg Cys Thr Ser
 165 170 175

Ile Ser Trp Glu Leu Arg Gln Thr Leu Ser Val Val Phe Asp Ala Phe
 180 185 190

Ile Thr Gly Gln Gly Lys Lys Asp Trp Ser Leu Phe Arg Met Phe Ser
 195 200 205

Arg Thr Leu Thr Glu Pro Cys Pro Leu Ala Ser Glu Ser Arg Val Tyr
 210 215 220

Val Asp Ile Thr Thr Tyr Asn Gln Asp Asn Glu Thr Leu Glu Val His
 225 230 235 240

Pro Pro Pro Thr Thr Thr Tyr Gln Asp Val Ile Leu Gly Thr Arg Lys
 245 250 255

Thr Tyr Ala Ile Tyr Asp Leu Leu Asp Thr Ala Met Ile Asn Asn Ser
 260 265 270

Arg Asn Leu Asn Ile Gln Leu Lys Trp Lys Arg Pro Pro Glu Asn Glu
 275 280 285

155

Ala Pro Pro Val Pro Phe Leu His Ala Gln Arg Tyr Val Ser Gly Tyr
 290 295 300

Gly Leu Gln Lys Gly Glu Leu Ser Thr Leu Leu Tyr Asn Thr His Pro
 305 310 315 320

Tyr Arg Ala Phe Pro Val Leu Leu Leu Asp Thr Val Pro Trp Tyr Leu
 325 330 335

Arg Leu Tyr Val His Thr Leu Thr Ile Thr Ser Lys Gly Lys Glu Asn
 340 345 350

Lys Pro Ser Tyr Ile His Tyr Gln Pro Ala Gln Asp Arg Leu Gln Pro
 355 360 365

His Leu Leu Glu Met Leu Ile Gln Leu Pro Ala Asn Ser Val Thr Lys
 370 375 380

Val Ser Ile Gln Phe Glu Arg Ala Leu Leu Lys Trp Thr Glu Tyr Thr
 385 390 395 400

Pro Asp Pro Asn His Gly Phe Tyr Val Ser Pro Ser Val Leu Ser Ala
 405 410 415

Leu Val Pro Ser Met Val Ala Ala Lys Pro Val Asp Trp Glu Glu Ser
 420 425 430

Pro Leu Phe Asn Ser Leu Phe Pro Val Ser Asp Gly Ser Asn Tyr Phe
 435 440 445

Val Arg Leu Tyr Thr Glu Pro Leu Leu Val Asn Leu Pro Thr Pro Asp
 450 455 460

Phe Ser Met Pro Tyr Asn Val Ile Cys Leu Thr Cys Thr Val Val Ala
 465 470 475 480

Val Cys Tyr Gly Ser Phe Tyr Asn Leu Leu Thr Arg Thr Phe Pro His
 485 490 495

Arg Gly Ala Pro His Arg Trp Pro Gly Gln Ala Ala Gly Gln Pro Tyr
 500 505 510

Pro Ala Arg Pro Ser Val Pro Pro Thr Leu Ile Leu Ala Leu Ser Ser
 515 520 525

Ser Cys Ser Cys Arg Phe Ser Leu Gly Arg Gly Ala Gln Gly Leu Phe
 530 535 540

Leu Pro Leu Ala Leu Leu Arg Val Gly Phe Xaa
 545 550 555

<210> 249

<211> 21

<212> PRT

<213> Homo sapiens

<400> 249

Thr Arg Pro Glu Lys Val Gln Ala Pro Leu Lys Trp Phe Lys Phe Gln

1

5

10

15

156

Ile Leu Asp Pro Pro
20

<210> 250

<211> 272

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (229)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 250

Ser Ala Glu Phe Gly Val Ala Pro Leu Pro Gly Arg Arg Gly Ser Pro
1 5 10 15

Val Arg Gln Leu Ala Gln Phe Arg Arg Arg Leu Leu Arg Gly Ser Gly
20 25 30

Gly Arg Gly Ala Pro Gly Arg Pro Pro Arg Cys Pro Gly Glu Ala Arg
35 40 45

Val Met Xaa Pro Pro Ser Cys Ile Gln Asp Glu Pro Phe Pro His Pro
50 55 60

Leu Glu Pro Glu Pro Gly Val Ser Ala Gln Pro Gly Pro Gly Lys Pro
65 70 75 80

Ser Asp Lys Arg Phe Arg Leu Trp Tyr Val Gly Gly Ser Cys Leu Asp
85 90 95

His Arg Thr Thr Leu Pro Met Leu Pro Trp Leu Met Ala Glu Ile Arg
100 105 110

Arg Arg Ser Gln Lys Pro Glu Ala Gly Gly Cys Gly Ala Pro Ala Ala
115 120 125

Arg Glu Val Ile Leu Val Leu Ser Ala Pro Phe Leu Arg Cys Val Pro
130 135 140

Ala Pro Gly Ala Gly Ala Ser Gly Gly Thr Ser Pro Ser Ala Thr Gln
145 150 155 160

Pro Asn Pro Ala Val Phe Ile Phe Glu His Lys Ala Gln His Ile Ser
165 170 175

Arg Phe Ile His Asn Ser His Asp Leu Thr Tyr Phe Ala Tyr Leu Ile
180 185 190

Lys Ala Gln Pro Asp Asp Pro Glu Ser Gln Met Ala Cys His Val Phe
195 200 205

157

Arg Ala Thr Asp Pro Ser Gln Val Pro Asp Val Ile Ser Ser Ile Arg
210 215 220

Gln Leu Ser Lys Xaa Ala Met Lys Glu Asp Ala Lys Pro Ser Lys Asp
225 230 235 240

Asn Glu Asp Ala Phe Tyr Asn Ser Gln Lys Phe Glu Val Leu Tyr Cys
245 250 255

Gly Lys Val Thr Val Thr Pro Gln Glu Gly Pro Leu Lys Pro His Arg
260 265 270

<210> 251

<211> 14

<212> PRT

<213> Homo sapiens

<400> 251

Pro Met Leu Pro Trp Leu Met Ala Glu Ile Arg Arg Arg Ser
1 5 10

<210> 252

<211> 19

<212> PRT

<213> Homo sapiens

<400> 252

Ile His Asn Ser His Asp Leu Thr Tyr Phe Ala Tyr Leu Ile Lys Ala
1 5 10 15

Gln Pro Asp

<210> 253

<211> 12

<212> PRT

<213> Homo sapiens

<400> 253

Lys Phe Glu Val Leu Tyr Cys Gly Lys Val Thr Val
1 5 10

<210> 254

<211> 13

<212> PRT

<213> Homo sapiens

<400> 254

Ile Ser Ser Ile Arg Gln Leu Ser Lys Ala Met Lys Glu
1 5 10

<210> 255

<211> 20

<212> PRT

<213> Homo sapiens

158

<400> 255

Gly Glu Arg Arg Asn Trp Gly Gly Glu Val Tyr Tyr Ser Thr Gly Tyr
 1 5 10 15

Ser Ser Arg Lys
 20

<210> 256

<211> 9

<212> PRT

<213> Homo sapiens

<400> 256

Glu Pro Gly Ala Ala Gln Glu Ser Trp
 1 5

<210> 257

<211> 202

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (108)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (120)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (138)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (165)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 257

Leu Cys Ala Arg Pro Ser Cys Ser Tyr Thr Gly Ala Glu Asn Gln Gly
 1 5 10 15

Gln Pro Arg Ser Pro Gly Trp Gly Ser Ser His Val Gly Trp Gly Trp
 20 25 30

Gly Val Gly Ser Pro Phe Leu Gly Ser Gln Glu Trp Ser Gly Leu Ala
 35 40 45

Pro Asp Leu Pro Asp Gln Glu Glu Glu Gln Pro Val Gly Arg His Ser
 50 55 60

Cys Pro Asp Met Ser Gln Cys Ile Lys Arg Gly His Gln Pro Val Gly
 65 70 75 80

Phe Ser Lys His Ala Trp Arg Cys Leu Val Gly Cys Cys Pro Trp Glu
 85 90 95

159

Glu Glu Lys Arg Ser Cys His Pro Phe Gly Ala Xaa Leu Leu Trp Val
 100 105 110
 Leu Arg Phe Ala Leu Gln Pro Xaa Val Tyr Glu Asp Pro Ala Ala Leu
 115 120 125
 Asp Gly Gly Glu Glu Gly Met Asp Ile Xaa Thr His Ile Leu Ala Leu
 130 135 140
 Ala Pro Arg Leu Leu Lys Asp Ser Gly Ser Ile Phe Leu Glu Val Asp
 145 150 155 160
 Pro Arg His Pro Xaa Leu Val Ser Ser Trp Leu Gln Ser Arg Pro Asp
 165 170 175
 Leu Tyr Leu Asn Leu Val Ala Val Arg Arg Asp Phe Cys Gly Arg Pro
 180 185 190
 Arg Phe Leu His Ile Arg Arg Ser Gly Pro
 195 200

<210> 258
 <211> 37
 <212> PRT
 <213> Homo sapiens

<400> 258
 Leu Cys Ala Arg Pro Ser Cys Ser Tyr Thr Gly Ala Glu Asn Gln Gly
 1 5 10 15
 Gln Pro Arg Ser Pro Gly Trp Gly Ser Ser His Val Gly Trp Gly Trp
 20 25 30
 Gly Val Gly Ser Pro
 35

<210> 259
 <211> 37
 <212> PRT
 <213> Homo sapiens

<400> 259
 Phe Leu Gly Ser Gln Glu Trp Ser Gly Leu Ala Pro Asp Leu Pro Asp
 1 5 10 15
 Gln Glu Glu Glu Gln Pro Val Gly Arg His Ser Cys Pro Asp Met Ser
 20 25 30
 Gln Cys Ile Lys Arg
 35

<210> 260
 <211> 37
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE

160

<222> (34)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 260

Gly	His	Gln	Pro	Val	Gly	Phe	Ser	Lys	His	Ala	Trp	Arg	Cys	Leu	Val
1				5					10					15	

Gly	Cys	Cys	Pro	Trp	Glu	Glu	Glu	Lys	Arg	Ser	Cys	His	Pro	Phe	Gly
			20					25					30		

Ala	Xaa	Leu	Leu	Trp
			35	

<210> 261

<211> 37

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (9)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (27)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 261

Val	Leu	Arg	Phe	Ala	Leu	Gln	Pro	Xaa	Val	Tyr	Glu	Asp	Pro	Ala	Ala
1				5					10					15	

Leu	Asp	Gly	Gly	Glu	Glu	Gly	Met	Asp	Ile	Xaa	Thr	His	Ile	Leu	Ala
			20					25					30		

Leu	Ala	Pro	Arg	Leu
				35

<210> 262

<211> 54

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (17)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 262

Leu	Lys	Asp	Ser	Gly	Ser	Ile	Phe	Leu	Glu	Val	Asp	Pro	Arg	His	Pro
1				5					10					15	

Xaa	Leu	Val	Ser	Ser	Trp	Leu	Gln	Ser	Arg	Pro	Asp	Leu	Tyr	Leu	Asn
			20					25					30		

Leu	Val	Ala	Val	Arg	Arg	Asp	Phe	Cys	Gly	Arg	Pro	Arg	Phe	Leu	His
		35					40						45		

Ile	Arg	Arg	Ser	Gly	Pro
-----	-----	-----	-----	-----	-----

161

50

<210> 263
<211> 19
<212> PRT
<213> Homo sapiens

<400> 263
Gln Glu Leu Leu Val Lys Ile Pro Leu Asp Met Val Ala Gly Phe Asn
1 5 10 15

Thr Pro Leu

<210> 264
<211> 26
<212> PRT
<213> Homo sapiens

<400> 264
Leu Arg Ile Gln Leu Leu His Lys Leu Ser Phe Leu Val Asn Ala Leu
1 5 10 15

Ala Lys Gln Val Met Asn Leu Leu Val Pro
20 25

<210> 265
<211> 20
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (2)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (10)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 265
His Xaa Ile Trp Leu Lys Val Ile Thr Xaa Asn Ile Leu Gln Leu Gln
1 5 10 15

Val Lys Pro Ser
20

<210> 266
<211> 58
<212> PRT
<213> Homo sapiens

<400> 266
Ala Gly Pro Trp Thr Phe Thr Leu Leu Cys Gly Leu Leu Ala Ala Thr
1 5 10 15

Leu Ile Gln Ala Thr Leu Ser Pro Thr Ala Val Leu Ile Leu Gly Pro
20 25 30

162

Lys Val Ile Lys Glu Lys Leu Thr Gln Glu Leu Lys Asp His Asn Ala
 35 40 45

Thr Ser Ile Leu Gln Gln Leu Pro Leu Leu
 50 55

<210> 267

<211> 15

<212> PRT

<213> Homo sapiens

<400> 267

His Phe Ile Ile Thr Leu Thr Thr Phe Phe Thr Asn Tyr Phe Leu
 1 5 10 15

<210> 268

<211> 99

<212> PRT

<213> Homo sapiens

<400> 268

Met Lys Ile Thr Phe Gln Asp Leu Phe Pro Met Trp Asn Ser Phe Lys
 1 5 10 15

Cys Phe Leu His Gly Asn Val Phe Ser Leu Phe Val Leu Phe Pro Leu
 20 25 30

Leu Thr Cys Phe Ser Phe Pro Tyr Thr Val Asn Ser Gly Thr Lys Leu
 35 40 45

Asp Trp Val Gly Trp Leu Val Gly Trp Phe Phe Leu Glu Phe Met Tyr
 50 55 60

Ile Asn Lys Gly Phe Glu Val Thr Ser Glu Asn Asn Ile Ser Lys Arg
 65 70 75 80

Val Leu Val Arg Glu Asn Ile Arg Ile Lys Ser Ser Pro Glu Arg Val
 85 90 95

Leu Arg Met

<210> 269

<211> 19

<212> PRT

<213> Homo sapiens

<400> 269

Arg Phe Trp Gly Ser Tyr Glu Pro His Phe Ser Gln Glu Val Ser Val
 1 5 10 15

Ile Pro Pro

<210> 270

<211> 56

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (32)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 270

Ile Arg Gly Asn Tyr Phe Ser Gly Arg Lys Lys Ser Ser Ser Asp Thr
 1 5 10 15

Pro Lys Gly Ser Lys Asp Lys Ile Ser Val Trp Asn Arg Ser Gln Xaa
 20 25 30

Ala Cys Ile Arg Ile Cys Lys Val His Pro Asn Tyr Ile Gln Ile Tyr
 35 40 45

Leu Trp His Ser Ala Thr Ser Phe
 50 55

<210> 271

<211> 74

<212> PRT

<213> Homo sapiens

<400> 271

Ala Gly Asn Gln Val Glu Pro Phe His Val Ser Leu Pro Ser Cys Leu
 1 5 10 15

Ser Pro Leu Pro His Leu Gly His Ser Met Gly Val Pro Ser Pro Thr
 20 25 30

Ala Trp Pro Ser Leu Ala Ser Phe His Thr Gln Lys Lys Ala Arg Ile
 35 40 45

Arg Gln Glu Glu Glu Ser Pro Pro Leu Pro Ser Pro Gln Glu Leu Ala
 50 55 60

Phe Ser Ala Leu Arg Val Phe Phe Arg Val
 65 70

<210> 272

<211> 38

<212> PRT

<213> Homo sapiens

<400> 272

Phe Ile Gln Gln Asn Ile Ser Phe Leu Leu Gly Tyr Ser Ile Pro Val
 1 5 10 15

Gly Cys Val Gly Leu Ala Phe Phe Ile Phe Leu Phe Ala Thr Pro Val
 20 25 30

Phe Ile Thr Lys Pro Pro
 35

<210> 273

<211> 347

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (16)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (340)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (341)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 273

Val	Ser	Ala	His	His	Pro	Ser	Gly	Ala	Asp	Glu	Gly	Val	Thr	Ala	Xaa
1				5					10					15	

Gln	Ile	Leu	Pro	Thr	Glu	Glu	Tyr	Glu	Glu	Ala	Met	Ser	Thr	Met	Gln
		20					25						30		

Val	Ser	Gln	Leu	Asp	Leu	Phe	Arg	Leu	Leu	Asp	Gln	Asn	Arg	Asp	Gly
		35					40					45			

His	Leu	Gln	Leu	Arg	Glu	Val	Leu	Ala	Gln	Thr	Arg	Leu	Gly	Asn	Gly
	50					55					60				

Trp	Trp	Met	Thr	Pro	Glu	Ser	Ile	Gln	Glu	Met	Tyr	Ala	Ala	Ile	Lys
65					70					75				80	

Ala	Asp	Pro	Asp	Gly	Asp	Gly	Val	Leu	Ser	Leu	Gln	Glu	Phe	Ser	Asn
				85					90					95	

Met	Asp	Leu	Arg	Asp	Phe	His	Lys	Tyr	Met	Arg	Ser	His	Lys	Ala	Glu
		100						105					110		

Ser	Ser	Glu	Leu	Val	Arg	Asn	Ser	His	His	Thr	Trp	Leu	Tyr	Gln	Gly
		115					120					125			

Glu	Gly	Ala	His	His	Ile	Met	Arg	Ala	Ile	Arg	Gln	Arg	Val	Leu	Arg
	130					135					140				

Leu	Thr	Arg	Leu	Ser	Pro	Glu	Ile	Val	Glu	Leu	Ser	Glu	Pro	Leu	Gln
145					150					155				160	

Val	Val	Arg	Tyr	Gly	Glu	Gly	Gly	His	Tyr	His	Ala	His	Val	Asp	Ser
				165					170					175	

Gly	Pro	Val	Tyr	Pro	Glu	Thr	Ile	Cys	Ser	His	Thr	Lys	Leu	Val	Ala
		180						185					190		

Asn	Glu	Ser	Val	Pro	Phe	Glu	Thr	Ser	Cys	Arg	Tyr	Met	Thr	Val	Leu
		195					200					205			

Phe	Tyr	Leu	Asn	Asn	Val	Thr	Gly	Gly	Gly	Glu	Thr	Val	Phe	Pro	Val
	210						215					220			

165
 Ala Asp Asn Arg Thr Tyr Asp Glu Met Ser Leu Ile Gln Asp Asp Val
 225 230 235 240
 Asp Leu Arg Asp Thr Arg Arg His Cys Asp Lys Gly Asn Leu Arg Val
 245 250 255
 Lys Pro Gln Gln Gly Thr Ala Val Phe Trp Tyr Asn Tyr Leu Pro Asp
 260 265 270
 Gly Gln Gly Trp Val Gly Asp Val Asp Asp Tyr Ser Leu His Gly Gly
 275 280 285
 Cys Leu Val Thr Arg Gly Thr Lys Trp Ile Ala Asn Asn Trp Ile Asn
 290 295 300
 Val Asp Pro Ser Arg Ala Arg Gln Ala Leu Phe Gln Gln Glu Met Ala
 305 310 315 320
 Arg Leu Ala Arg Glu Gly Gly Thr Asp Ser Gln Pro Glu Trp Ala Leu
 325 330 335
 Asp Arg Ala Xaa Xaa Asp Ala Arg Val Glu Leu
 340 345

<210> 274
 <211> 6
 <212> PRT
 <213> Homo sapiens

<400> 274
 Ala Val Phe Trp Tyr Asn
 1 5

<210> 275
 <211> 18
 <212> PRT
 <213> Homo sapiens

<400> 275
 Thr Val Leu Phe Tyr Leu Asn Asn Val Thr Gly Gly Gly Glu Thr Val
 1 5 10 15

Phe Pro

<210> 276
 <211> 59
 <212> PRT
 <213> Homo sapiens

<400> 276
 Asp Leu Phe Arg Leu Leu Asp Gln Asn Arg Asp Gly His Leu Gln Leu
 1 5 10 15

Arg Glu Val Leu Ala Gln Thr Arg Leu Gly Asn Gly Trp Trp Met Thr
 20 25 30

Pro Glu Ser Ile Gln Glu Met Tyr Ala Ala Ile Lys Ala Asp Pro Asp
 35 40 45

166

Gly Asp Gly Val Leu Ser Leu Gln Glu Phe Ser
 50 55

<210> 277

<211> 38

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (16)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 277

Val Ser Ala His His Pro Ser Gly Ala Asp Glu Gly Val Thr Ala Xaa
 1 5 10 15

Gln Ile Leu Pro Thr Glu Glu Tyr Glu Glu Ala Met Ser Thr Met Gln
 20 25 30

Val Ser Gln Leu Asp Leu
 35

<210> 278

<211> 38

<212> PRT

<213> Homo sapiens

<400> 278

Phe Arg Leu Leu Asp Gln Asn Arg Asp Gly His Leu Gln Leu Arg Glu
 1 5 10 15

Val Leu Ala Gln Thr Arg Leu Gly Asn Gly Trp Trp Met Thr Pro Glu
 20 25 30

Ser Ile Gln Glu Met Tyr
 35

<210> 279

<211> 38

<212> PRT

<213> Homo sapiens

<400> 279

Ala Ala Ile Lys Ala Asp Pro Asp Gly Asp Gly Val Leu Ser Leu Gln
 1 5 10 15

Glu Phe Ser Asn Met Asp Leu Arg Asp Phe His Lys Tyr Met Arg Ser
 20 25 30

His Lys Ala Glu Ser Ser
 35

<210> 280

<211> 38

<212> PRT

<213> Homo sapiens

167

<400> 280

Glu Leu Val Arg Asn Ser His His Thr Trp Leu Tyr Gln Gly Glu Gly
1 5 10 15

Ala His His Ile Met Arg Ala Ile Arg Gln Arg Val Leu Arg Leu Thr
20 25 30

Arg Leu Ser Pro Glu Ile
35

<210> 281

<211> 38

<212> PRT

<213> Homo sapiens

<400> 281

Val Glu Leu Ser Glu Pro Leu Gln Val Val Arg Tyr Gly Glu Gly Gly
1 5 10 15

His Tyr His Ala His Val Asp Ser Gly Pro Val Tyr Pro Glu Thr Ile
20 25 30

Cys Ser His Thr Lys Leu
35

<210> 282

<211> 38

<212> PRT

<213> Homo sapiens

<400> 282

Val Ala Asn Glu Ser Val Pro Phe Glu Thr Ser Cys Arg Tyr Met Thr
1 5 10 15

Val Leu Phe Tyr Leu Asn Asn Val Thr Gly Gly Gly Glu Thr Val Phe
20 25 30

Pro Val Ala Asp Asn Arg
35

<210> 283

<211> 38

<212> PRT

<213> Homo sapiens

<400> 283

Thr Tyr Asp Glu Met Ser Leu Ile Gln Asp Asp Val Asp Leu Arg Asp
1 5 10 15

Thr Arg Arg His Cys Asp Lys Gly Asn Leu Arg Val Lys Pro Gln Gln
20 25 30

Gly Thr Ala Val Phe Trp
35

<210> 284

<211> 38

<212> PRT

<213> Homo sapiens

168

<400> 284

Tyr Asn Tyr Leu Pro Asp Gly Gln Gly Trp Val Gly Asp Val Asp Asp
 1 5 10 15

Tyr Ser Leu His Gly Gly Cys Leu Val Thr Arg Gly Thr Lys Trp Ile
 20 25 30

Ala Asn Asn Trp Ile Asn
 35

<210> 285

<211> 43

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (36)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (37)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 285

Val Asp Pro Ser Arg Ala Arg Gln Ala Leu Phe Gln Gln Glu Met Ala
 1 5 10 15

Arg Leu Ala Arg Glu Gly Gly Thr Asp Ser Gln Pro Glu Trp Ala Leu
 20 25 30

Asp Arg Ala Xaa Xaa Asp Ala Arg Val Glu Leu
 35 40

<210> 286

<211> 15

<212> PRT

<213> Homo sapiens

<400> 286

Leu Leu Ala Asp Leu Met Arg Asn Tyr Asp Pro His Leu Arg Pro
 1 5 10 15

<210> 287

<211> 19

<212> PRT

<213> Homo sapiens

<400> 287

Ile Ser Val Thr Tyr Phe Pro Phe Asp Trp Gln Asn Cys Ser Leu Ile
 1 5 10 15

Phe Gln Ser

<210> 288

<211> 16

169

<212> PRT

<213> Homo sapiens

<400> 288

Ser Met Ala Arg Gly Val Arg Lys Val Phe Leu Arg Leu Leu Pro Gln
1 5 10 15

<210> 289

<211> 18

<212> PRT

<213> Homo sapiens

<400> 289

Gln Ala Ser Pro Ala Ile Gln Ala Cys Val Asp Ala Cys Asn Leu Met
1 5 10 15

Ala Arg

<210> 290

<211> 17

<212> PRT

<213> Homo sapiens

<400> 290

Tyr Asn Gln Val Pro Asp Leu Pro Phe Pro Gly Asp Pro Arg Pro Tyr
1 5 10 15

Leu

<210> 291

<211> 15

<212> PRT

<213> Homo sapiens

<400> 291

Cys Ser Ile Ser Val Thr Tyr Phe Pro Phe Asp Trp Gln Asn Cys
1 5 10 15

<210> 292

<211> 18

<212> PRT

<213> Homo sapiens

<400> 292

Val Leu Lys Tyr Ala Leu Phe Leu Val Leu Lys Asn Tyr Tyr Tyr Cys
1 5 10 15

Pro Tyr

<210> 293

<211> 315

<212> PRT

<213> Homo sapiens

170

<400> 293

Met Arg Glu Tyr Gly Val Glu Arg Asp Leu Ala Val Tyr Asn Gln Leu
 1 5 10 15

Leu Asn Ile Phe Pro Lys Glu Val Phe Arg Pro Arg Asn Ile Ile Gln
 20 25 30

Arg Ile Phe Val His Tyr Pro Arg Gln Gln Glu Cys Gly Ile Ala Val
 35 40 45

Leu Glu Gln Met Glu Asn His Gly Val Met Pro Asn Lys Glu Thr Glu
 50 55 60

Phe Leu Leu Ile Gln Ile Phe Gly Arg Lys Ser Tyr Pro Met Leu Lys
 65 70 75 80

Leu Val Arg Leu Lys Leu Trp Phe Pro Arg Phe Met Asn Val Asn Pro
 85 90 95

Phe Pro Val Pro Arg Asp Leu Pro Gln Asp Pro Val Glu Leu Ala Met
 100 105 110

Phe Gly Leu Arg His Met Glu Pro Asp Leu Ser Ala Arg Val Thr Ile
 115 120 125

Tyr Gln Val Pro Leu Pro Lys Asp Ser Thr Gly Ala Ala Asp Pro Pro
 130 135 140

Gln Pro His Ile Val Gly Ile Gln Ser Pro Asp Gln Gln Ala Ala Leu
 145 150 155 160

Ala Arg His Asn Pro Ala Arg Pro Val Phe Val Glu Gly Pro Phe Ser
 165 170 175

Leu Trp Leu Arg Asn Lys Cys Val Tyr Tyr His Ile Leu Arg Ala Asp
 180 185 190

Leu Leu Pro Pro Glu Glu Arg Glu Val Glu Glu Thr Pro Glu Glu Trp
 195 200 205

Asn Leu Tyr Tyr Pro Met Gln Leu Asp Leu Glu Tyr Val Arg Ser Gly
 210 215 220

Trp Asp Asn Tyr Glu Phe Asp Ile Asn Glu Val Glu Glu Gly Pro Val
 225 230 235 240

Phe Ala Met Cys Met Ala Gly Ala His Asp Gln Ala Thr Met Ala Lys
 245 250 255

Trp Ile Gln Gly Leu Gln Glu Thr Asn Pro Thr Leu Ala Gln Ile Pro
 260 265 270

Val Val Phe Arg Leu Ala Gly Ser Thr Arg Glu Leu Gln Thr Ser Ser
 275 280 285

Ala Gly Leu Glu Glu Pro Pro Leu Pro Glu Asp His Gln Glu Glu Asp
 290 295 300

171
Asp Asn Leu Gln Arg Gln Gln Gln Gly Gln Ser
305 310 315

<210> 294
<211> 19
<212> PRT
<213> Homo sapiens

<400> 294
Phe Gln Phe Gly Trp Ala Ser Thr Gln Ile Ser His Leu Ser Leu Ile
1 5 10 15

Pro Glu Leu

<210> 295
<211> 14
<212> PRT
<213> Homo sapiens

<400> 295
Leu Arg Tyr Ala Phe Thr Val Val Ala Asn Ile Thr Val Tyr
1 5 10

<210> 296
<211> 17
<212> PRT
<213> Homo sapiens

<400> 296
Phe Val Tyr Gly Ser Met Ser Phe Leu Asp Lys Val Ala Asn Gly Leu
1 5 10 15

Ala

<210> 297
<211> 17
<212> PRT
<213> Homo sapiens

<400> 297
Trp His Leu Val Gly Thr Val Cys Val Leu Leu Ser Phe Pro Phe Ile
1 5 10 15

Phe

<210> 298
<211> 15
<212> PRT
<213> Homo sapiens

<400> 298
Gly His Phe Leu Asn Asp Leu Cys Ala Ser Met Trp Phe Thr Tyr
1 5 10 15

<210> 299
<211> 40

172

<212> PRT

<213> Homo sapiens

<400> 299

Ala Ile Pro Leu Arg Val Leu Val Val Leu Trp Ala Phe Val Leu Gly
 1 5 10 15

Leu Ser Arg Val Met Leu Gly Arg His Asn Val Thr Asp Val Ala Phe
 20 25 30

Gly Phe Phe Leu Gly Tyr Met Gln
 35 40

<210> 300

<211> 13

<212> PRT

<213> Homo sapiens

<400> 300

Val Gly Leu Ser Arg Val Leu Gly Arg His Thr Asp Val
 1 5 10

<210> 301

<211> 17

<212> PRT

<213> Homo sapiens

<400> 301

Ser Phe Tyr Lys Met Lys Arg Asn Ser Tyr Asp Arg Leu Arg Lys Val
 1 5 10 15

Val

<210> 302

<211> 39

<212> PRT

<213> Homo sapiens

<400> 302

Leu His Gln Leu Arg Pro Pro His Arg Phe Pro Leu Ile Pro Pro Ala
 1 5 10 15

Ala Ala Glu Gly Ala Gly Ala Pro Pro Gly Cys Gly Tyr Cys Val Phe
 20 25 30

Trp Leu Leu Asn Pro Leu Pro
 35

<210> 303

<211> 72

<212> PRT

<213> Homo sapiens

<400> 303

Met Pro Trp Lys Arg Ala Val Val Leu Leu Met Leu Trp Phe Ile Gly
 1 5 10 15

Gln Ala Met Trp Leu Ala Pro Ala Tyr Val Leu Glu Phe Gln Gly Lys

Leu Xaa Arg Val Pro Gly Glu Ala Glu Ser Leu Cys Ala Leu Ser Pro

174

50	55	60
Gly Ser Ala Leu Arg Phe Pro Ala Ala Ser Cys Ser Arg Pro Xaa Arg		
65	70	75
Glu Pro Ser Gly Asp Glu Gly Thr Ala Gly Ala Leu Pro Ser Pro Trp		
85	90	95
Leu Ala Ala Leu Gly Pro Gly Gly Arg Pro Ala Val Arg Arg Val Leu		
100	105	110
Pro Arg Leu Gly Gly Arg Ala Gly Gln Leu Pro Arg Gly Leu Pro Val		
115	120	125
Pro Arg Gly Leu Arg His Ala Gly Arg Tyr His Leu Leu Arg Leu Leu		
130	135	140
Arg Ala Pro Leu Leu Leu Arg Arg Gly Arg Arg Gln Ala Gly Ala Gly		
145	150	155
Arg Leu His Gln Arg Pro Pro Arg Thr Gly Ala Pro Arg His His Cys		
165	170	175
Ala Ala Cys Leu Arg Pro Leu Ser His Arg Arg Leu His Leu His Cys		
180	185	190
Val His His Pro Gly Leu Cys Ser Gly Tyr Leu Leu Leu His Leu Phe		
195	200	205
Glu Thr Gln Gly Ala Leu Ala Ala Ala Asn Pro Leu Leu Thr Pro Gln		
210	215	220
Leu Ser Asp Arg Asp Pro Ala His Asp Pro Asp Leu His Gln Pro Gln		
225	230	235
Gly Thr Leu Pro Ala Val Gln His Ser His Glu Leu Gln Leu His Arg		
245	250	255
Arg Leu His Pro Gln Val Leu Leu Ser His Leu Val Ser Trp Cys His		
260	265	270
Pro Ser Ile Ser Leu Thr Pro Phe Ser Arg Ser Pro His Trp Leu Gly		
275	280	285
Arg Ala Val Gln Thr Phe Ser Ser Xaa		
290	295	

<210> 306

<211> 38

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (12)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 306

Ala Gly Leu Pro Gly Ala Leu Thr Ala Pro Ala Xaa His His His Ala

175
1 5 10 15
Asp Ser Arg Pro Ala Glu Leu Val Val Gln Pro Leu Ser Pro Pro Arg
20 25 30

Pro Leu Leu Ser His Ala
35

<210> 307
<211> 40
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (12)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 307
Gly Leu Ala Ser Ala Ala Gly Ala Ser Ser Leu Xaa Arg Val Pro Gly
1 5 10 15

Glu Ala Glu Ser Leu Cys Ala Leu Ser Pro Gly Ser Ala Leu Arg Phe
20 25 30

Pro Ala Ala Ser Cys Ser Arg Pro
35 40

<210> 308
<211> 40
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (1)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 308
Xaa Arg Glu Pro Ser Gly Asp Glu Gly Thr Ala Gly Ala Leu Pro Ser
1 5 10 15

Pro Trp Leu Ala Ala Leu Gly Pro Gly Gly Arg Pro Ala Val Arg Arg
20 25 30

Val Leu Pro Arg Leu Gly Gly Arg
35 40

<210> 309
<211> 40
<212> PRT
<213> Homo sapiens

<400> 309
Ala Gly Gln Leu Pro Arg Gly Leu Pro Val Pro Arg Gly Leu Arg His
1 5 10 15

Ala Gly Arg Tyr His Leu Leu Arg Leu Leu Arg Ala Pro Leu Leu Leu
20 25 30

176

Arg Arg Gly Arg Arg Gln Ala Gly
 35 40

<210> 310

<211> 40

<212> PRT

<213> Homo sapiens

<400> 310

Ala Gly Arg Leu His Gln Arg Pro Pro Arg Thr Gly Ala Pro Arg His
 1 5 10 15

His Cys Ala Ala Cys Leu Arg Pro Leu Ser His Arg Arg Leu His Leu
 20 25 30

His Cys Val His His Pro Gly Leu
 35 40

<210> 311

<211> 40

<212> PRT

<213> Homo sapiens

<400> 311

Cys Ser Gly Tyr Leu Leu Leu His Leu Phe Glu Thr Gln Gly Ala Leu
 1 5 10 15

Ala Ala Ala Asn Pro Leu Leu Thr Pro Gln Leu Ser Asp Arg Asp Pro
 20 25 30

Ala His Asp Pro Asp Leu His Gln
 35 40

<210> 312

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 312

Pro Gln Gly Thr Leu Pro Ala Val Gln His Ser His Glu Leu Gln Leu
 1 5 10 15

His Arg Arg Leu His Pro Gln Val Leu Leu Ser His Leu Val Ser Trp
 20 25 30

Cys His Pro Ser Ile Ser Leu Thr Pro Phe Ser Arg Ser Pro His Trp
 35 40 45

Leu Gly Arg Ala Val Gln Thr Phe Ser Ser Xaa
 50 55

<210> 313

<211> 28

177

<212> PRT

<213> Homo sapiens

<400> 313

Val Ala His Thr Cys Asn Leu Ser Thr Leu Gly Gly Gln Gly Gly Arg
1 5 10 15

Ile Glu Arg Thr Ala Gly Gln Glu Phe Lys Thr Ser
20 25

<210> 314

<211> 19

<212> PRT

<213> Homo sapiens

<400> 314

Thr Ile Lys Met Gln Thr Glu Asn Leu Gly Val Val Tyr Tyr Val Asn
1 5 10 15

Lys Asp Phe

<210> 315

<211> 13

<212> PRT

<213> Homo sapiens

<400> 315

Val Glu Glu Asp Tyr Val Thr Asn Ile Arg Asn Asn Cys
1 5 10

<210> 316

<211> 7

<212> PRT

<213> Homo sapiens

<400> 316

Met Val Ser Asn Pro Pro Tyr
1 5

<210> 317

<211> 5

<212> PRT

<213> Homo sapiens

<400> 317

His Ala Ser Glu Leu
1 5

<210> 318

<211> 35

<212> PRT

<213> Homo sapiens

<400> 318

Leu Val Ala Leu Asp Arg Met Glu Tyr Val Arg Thr Phe Arg Lys Arg
1 5 10 15

Glu Asp Leu Arg Gly Arg Leu Phe Trp Val Ala Leu Asp Leu Leu Asp

20 25 178 30

Leu Leu Asp
35

<210> 319
<211> 88
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (21)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 319
Ser Val Ala Leu Phe Tyr Asn Phe Gly Lys Ser Trp Lys Ser Asp Pro
1 5 10 15
Gly Ile Ile Lys Xaa Thr Glu Glu Gln Lys Lys Lys Thr Ile Val Glu
20 25 30
Leu Ala Glu Thr Gly Ser Leu Asp Leu Ser Ile Phe Cys Ser Thr Cys
35 40 45
Leu Ile Arg Lys Pro Val Arg Ser Lys His Cys Gly Val Cys Asn Arg
50 55 60
Cys Ile Ala Lys Phe Asp His His Cys Pro Trp Val Gly Asn Cys Val
65 70 75 80
Gly Ala Gly Asn His Arg Tyr Phe
85

<210> 320
<211> 12
<212> PRT
<213> Homo sapiens

<400> 320
Phe Asp His His Cys Pro Trp Val Gly Asn Cys Val
1 5 10

<210> 321
<211> 20
<212> PRT
<213> Homo sapiens

<400> 321
Gln Met Tyr Gln Ile Ser Cys Leu Gly Ile Thr Thr Asn Glu Arg Met
1 5 10 15
Asn Ala Arg Arg
20

<210> 322
<211> 12
<212> PRT
<213> Homo sapiens

179

<400> 322

Arg Val Thr Ser Ser Leu Ala Met Leu Ser Asp Ser
 1 5 10

<210> 323

<211> 15

<212> PRT

<213> Homo sapiens

<400> 323

Ala Ile Glu Arg Phe Ile Glu Pro His Glu Met Gln Gln Pro Leu
 1 5 10 15

<210> 324

<211> 49

<212> PRT

<213> Homo sapiens

<400> 324

Asn Ala Leu Val Phe Tyr Phe Ser Trp Lys Gly Cys Ser Glu Gly Asp
 1 5 10 15

Phe Cys Val Asn Pro Cys Phe Pro Asp Pro Cys Lys Pro Phe Val Glu
 20 25 30

Ile Ile Asn Ser Thr His Ala Ser Val Tyr Glu Ala Gly Pro Cys Trp
 35 40 45

Val

<210> 325

<211> 307

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (148)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 325

Ala Gly Ile Arg His Glu Arg Asn Arg Gly Arg Leu Leu Cys Met Leu
 1 5 10 15

Ala Leu Thr Phe Met Phe Met Val Leu Glu Val Val Val Ser Arg Val
 20 25 30

Thr Ser Ser Leu Ala Met Leu Ser Asp Ser Phe His Met Leu Ser Asp
 35 40 45

Val Leu Ala Leu Val Val Ala Leu Val Ala Glu Arg Phe Ala Arg Arg
 50 55 60

Thr His Ala Thr Gln Lys Asn Thr Phe Gly Trp Ile Arg Ala Glu Val
 65 70 75 80

Met Gly Ala Leu Val Asn Ala Ile Phe Leu Thr Gly Leu Cys Phe Ala

[illegible]

20 25 181 30
 Asp His Val Ile Lys Ile Thr Arg Ile Glu Val Gly Asp Val Asn Pro
 35 40 45
 Ser Glu Thr Gln Tyr Ile Ser Glu Pro Lys Leu Cys Pro Glu Cys Arg
 50 55 60
 Glu Gly Leu Leu Cys Gln Gln Gln Arg Asp Leu Arg Glu Tyr Thr Gln
 65 70 75 80
 Ala Thr Ile Tyr Val His Lys Val Val Asp Asn Lys Lys Val Met Lys
 85 90 95
 Asp Ser Ala Pro Glu Leu Asn Val Ser Ser Ser Glu Thr Glu Glu Asp
 100 105 110
 Lys Glu Glu Ala Lys Pro Asp Gly Glu Lys Asp Pro Asp Phe Asn Gln
 115 120 125
 Ser Xaa Gly Gly Thr Lys Arg Gln Lys Ile Ser His Gln Asn Tyr Ile
 130 135 140
 Ala Tyr Gln Lys Gln Val Ile Arg Arg Ser Met Arg His Arg Lys Val
 145 150 155 160
 Arg Gly Glu Lys Ala Leu Leu Val Ser Ala Asn Gln Thr Leu Lys Glu
 165 170 175
 Leu Lys Ile Gln Ile Met His Ala Phe Ser Val Ala Pro Phe Asp Gln
 180 185 190
 Asn Leu Ser Ile Asp Gly Lys Ile Leu Ser Asp Asp Cys Ala Thr Leu
 195 200 205
 Gly Thr Leu Gly Val Ile Pro Glu Ser Val Ile Leu Leu Lys Ala Asp
 210 215 220
 Glu Pro Ile Ala Asp Tyr Ala Ala Met Asp Asp Val Met Gln Val Cys
 225 230 235 240
 Met Pro Glu Glu Gly Phe Lys Gly Thr Gly Leu Leu Gly His
 245 250

182

<213> Homo sapiens

<400> 328

Lys Glu Leu Lys Ile Gln Ile Met His Ala Phe Ser Val Ala Pro Phe
1 5 10 15

Asp Gln

<210> 329

<211> 58

<212> PRT

<213> Homo sapiens

<400> 329

Phe Gln Asp Lys Asn Arg Pro Cys Leu Ser Asn Trp Pro Glu Asp Thr
1 5 10 15

Asp Val Leu Tyr Ile Val Ser Gln Phe Phe Val Glu Glu Trp Arg Lys
20 25 30

Phe Val Arg Lys Pro Thr Arg Cys Ser Pro Val Ser Ser Val Gly Asn
35 40 45

Ser Ala Leu Leu Cys Pro His Gly Gly Leu
50 55

<210> 330

<211> 42

<212> PRT

<213> Homo sapiens

<400> 330

Met Phe Thr Phe Ala Ser Met Thr Lys Glu Asp Ser Lys Leu Ile Ala
1 5 10 15

Leu Ile Trp Pro Ser Glu Trp Gln Met Ile Gln Lys Leu Phe Val Val
20 25 30

Asp His Val Ile Lys Ile Thr Arg Ile Glu
35 40

<210> 331

<211> 42

<212> PRT

<213> Homo sapiens

<400> 331

Val Gly Asp Val Asn Pro Ser Glu Thr Gln Tyr Ile Ser Glu Pro Lys
1 5 10 15

Leu Cys Pro Glu Cys Arg Glu Gly Leu Leu Cys Gln Gln Gln Arg Asp
20 25 30

Leu Arg Glu Tyr Thr Gln Ala Thr Ile Tyr
35 40

<210> 332

<211> 42

183

<212> PRT

<213> Homo sapiens

<400> 332

Val His Lys Val Val Asp Asn Lys Lys Val Met Lys Asp Ser Ala Pro
 1 5 10 15

Glu Leu Asn Val Ser Ser Ser Glu Thr Glu Glu Asp Lys Glu Glu Ala
 20 25 30

Lys Pro Asp Gly Glu Lys Asp Pro Asp Phe
 35 40

<210> 333

<211> 42

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (4)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 333

Asn Gln Ser Xaa Gly Gly Thr Lys Arg Gln Lys Ile Ser His Gln Asn
 1 5 10 15

Tyr Ile Ala Tyr Gln Lys Gln Val Ile Arg Arg Ser Met Arg His Arg
 20 25 30

Lys Val Arg Gly Glu Lys Ala Leu Leu Val
 35 40

<210> 334

<211> 42

<212> PRT

<213> Homo sapiens

<400> 334

Ser Ala Asn Gln Thr Leu Lys Glu Leu Lys Ile Gln Ile Met His Ala
 1 5 10 15

Phe Ser Val Ala Pro Phe Asp Gln Asn Leu Ser Ile Asp Gly Lys Ile
 20 25 30

Leu Ser Asp Asp Cys Ala Thr Leu Gly Thr
 35 40

<210> 335

<211> 44

<212> PRT

<213> Homo sapiens

<400> 335

Leu Gly Val Ile Pro Glu Ser Val Ile Leu Leu Lys Ala Asp Glu Pro
 1 5 10 15

Ile Ala Asp Tyr Ala Ala Met Asp Asp Val Met Gln Val Cys Met Pro
 20 25 30

184

Glu Glu Gly Phe Lys Gly Thr Gly Leu Leu Gly His
 35 40

<210> 336
 <211> 18
 <212> PRT
 <213> Homo sapiens

<400> 336
 Arg Gly Glu Arg Ser Glu Glu Leu Leu Gly Arg Glu Gly Leu Ser Gly
 1 5 10 15

Ser Gln

<210> 337
 <211> 179
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (119)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (123)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (177)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 337
 Ala Glu Ala Ala Glu Gly Glu Lys Gly Val Arg Ser Cys Trp Ala Glu
 1 5 10 15

Arg Asp Cys Pro Ala Pro Arg Cys Trp Ala Ser Trp Gly Ala Gln Pro
 20 25 30

Ser Trp Asp Gly Ser Gln Val Leu Leu Trp Arg Ser Cys Cys Cys Cys
 35 40 45

Cys Cys Trp Pro Pro Ala Phe Ser Thr Asp Gly Arg Thr Val Thr Trp
 50 55 60

Arg Gly Thr Val Gln Leu Gln Gly Glu Thr Glu Ser Ala Gly Pro Ser
 65 70 75 80

Leu Gly Pro Ser Gly Gly Gly Ala Thr Trp Glu Ser Phe Thr Ile Thr
 85 90 95

Val Ile Leu Ala Thr Tyr Leu Met Cys Arg Met Trp Ala Ser Thr Thr
 100 105 110

Thr Thr Thr Pro Ala Thr Xaa Leu Thr Thr Xaa Thr Thr Thr Thr Thr

115 120 125 185

Pro Thr Ala Thr Ile Pro Ala Thr Leu Ala Glu Ala Ala Val Ala Gly
130 135 140

Ala Cys Gly Gln Gln Leu Pro Leu Pro Ser His Leu Phe Pro Gly Gln
145 150 155 160

Val Asp Pro Met Phe Pro Cys Gly Arg Met His Leu Trp Gly Glu Arg
165 170 175

Xaa Glu Gln

<210> 338
<211> 12
<212> PRT
<213> Homo sapiens

<400> 338
Phe His Gly Leu Gly Arg Leu His Thr Val His Leu
1 5 10

<210> 339
<211> 21
<212> PRT
<213> Homo sapiens

<400> 339
Ala Ala Phe Thr Gly Leu Ala Leu Leu Glu Gln Leu Asp Leu Ser Asp
1 5 10 15

Asn Ala Gln Leu Arg
20

<210> 340
<211> 9
<212> PRT
<213> Homo sapiens

<400> 340
Ala Phe Arg Gly Leu His Ser Leu Asp
1 5

<210> 341
<211> 13
<212> PRT
<213> Homo sapiens

<400> 341
His Glu Val Pro Asp Ala Pro Arg Pro Thr Pro Thr Xaa
1 5 10

<210> 342
<211> 101
<212> PRT
<213> Homo sapiens

<400> 342

186

Met Val Val Ala Asp Arg Asn Arg Ala Ser Ser Ser Ser Tyr Leu Cys
 1 5 10 15

Leu Leu Leu Phe Ser Leu Ser Leu Phe Leu Cys His Glu Thr Val Cys
 20 25 30

Asp Arg Ala Thr Cys Leu Phe Phe Phe Leu Lys Phe Phe Phe Leu Phe
 35 40 45

Met Cys Arg Cys Met Ser Trp Gly Phe Lys Asn Phe Lys Ala Gly Leu
 50 55 60

Leu Met Gln Ser Met Pro Thr Ser Gly Ile Leu Arg Glu Arg Lys Arg
 65 70 75 80

Leu His Val Val Arg Ile Pro Gln Gly Thr Glu Lys Lys Leu Glu Thr
 85 90 95

Val Glu Met Gln Ile
 100

<210> 343
 <211> 12
 <212> PRT
 <213> Homo sapiens

<400> 343
 Ile Pro Gln Gly Thr Glu Lys Lys Leu Glu Thr Val
 1 5 10

<210> 344
 <211> 37
 <212> PRT
 <213> Homo sapiens

<400> 344
 Asn Pro Arg Leu Pro Leu Pro Arg Gly Gly Ser Leu Arg Leu Leu Ser
 1 5 10 15

Ser Pro Ala Asn Ser Asn Asn Ala Lys Ala Tyr Pro Phe Ser Arg Phe
 20 25 30

Pro Ser Pro Ile Phe
 35

<210> 345
 <211> 48
 <212> PRT
 <213> Homo sapiens

<400> 345
 Met Val Gln Glu Ala Pro Ala Leu Val Arg Leu Ser Leu Gly Ser His
 1 5 10 15

Arg Val Lys Gly Pro Leu Pro Val Leu Lys Leu Gln Pro Glu Gly Trp
 20 25 30

Ser Pro Ser Thr Leu Trp Ser Cys Ala Ser Val Trp Lys Asp Ser Cys
 35 40 45

<210> 346
<211> 122
<212> PRT
<213> Homo sapiens

<400> 346
Ala Leu Ala Ser Ser Leu Val Ala Glu Asn Gln Gly Phe Val Ala Ala
1 5 10 15
Leu Met Val Gln Glu Ala Pro Ala Leu Val Arg Leu Ser Leu Gly Ser
20 25 30
His Arg Val Lys Gly Pro Leu Pro Val Leu Lys Leu Gln Pro Glu Gly
35 40 45
Trp Ser Pro Ser Thr Leu Trp Ser Cys Ala Ser Val Trp Lys Asp Ser
50 55 60
Cys Met His Pro Trp Arg Leu Ser Met Cys Pro Ala Cys Val Leu Ala
65 70 75 80
Ala Leu Pro Ala Leu Cys Ser Cys Leu Cys Ser Pro Asp Ala Arg Pro
85 90 95
Pro His Gly Trp Met Ser Met Pro Phe Thr Pro His Pro Leu Val Ser
100 105 110
Arg Ala Met Pro Thr Cys His Pro Cys Ser
115 120

<210> 347
<211> 33
<212> PRT
<213> Homo sapiens

<400> 347
Phe Tyr Phe Ile Thr Leu Ile Phe Phe Leu Ala Trp Leu Val Lys Asn
1 5 10 15
Val Phe Ile Ala Val Ile Ile Glu Thr Phe Ala Glu Ile Arg Val Gln
20 25 30

Phe

<210> 348
<211> 15
<212> PRT
<213> Homo sapiens

<400> 348
Ser Ile Phe Thr Val Tyr Glu Ala Ala Ser Gln Glu Gly Trp Val
1 5 10 15

<210> 349

188

<211> 21

<212> PRT

<213> Homo sapiens

<400> 349

His	Glu	Gly	Thr	Ser	Ile	Phe	Thr	Val	Tyr	Glu	Ala	Ala	Ser	Gln	Glu
1				5					10					15	

Gly	Trp	Val	Phe	Leu
				20

<210> 350

<211> 8

<212> PRT

<213> Homo sapiens

<400> 350

Cys	Lys	Thr	Ser	Phe	Gly	Leu	Ala
1				5			

<210> 351

<211> 122

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (73)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 351

Met	Ile	Thr	Leu	Ser	Ser	Ala	Phe	Ser	Ala	Lys	Gln	Lys	Thr	His	Ala
1				5					10					15	

His	Lys	Asn	Thr	His	Ala	Cys	Met	Cys	Ala	Thr	Asp	Met	Ala	Asn	Pro
			20					25						30	

Lys	Leu	Val	Leu	His	Phe	Glu	Val	Ile	Val	Ala	Leu	Leu	Ser	Leu	Leu
		35					40						45		

Gln	Thr	Ile	Leu	Ser	Leu	Leu	Leu	Gly	Gln	Arg	Thr	Trp	Leu	Ala	His
		50				55						60			

Leu	Tyr	Val	Leu	Ser	Thr	Glu	Asn	Xaa	Ala	Leu	His	Thr	Val	Gly	Thr
	65				70					75				80	

Gln	Lys	His	Leu	Leu	Pro	His	Asp	Trp	Cys	Phe	Gly	Lys	His	Cys	Val
				85					90					95	

Ser	Cys	Arg	His	His	Ile	Phe	His	Arg	Phe	Cys	Ser	Ile	Phe	Ser	Ser
			100					105						110	

Thr	Leu	Lys	Arg	Ser	Gln	Gly	Phe	Glu	Gly
		115						120	

<210> 352

<211> 13

<212> PRT

<213> Homo sapiens

189

<400> 352

Cys Ala Ala Pro Gly Asn Lys Thr Ser His Leu Ala Ala
1 5 10

<210> 353

<211> 24

<212> PRT

<213> Homo sapiens

<400> 353

Glu His Pro Leu Tyr Arg Ala Gly His Leu Ile Leu Gln Asp Arg Ala
1 5 10 15

Ser Cys Leu Pro Ala Met Leu Leu
20

<210> 354

<211> 15

<212> PRT

<213> Homo sapiens

<400> 354

Leu Leu Asp Pro Ser Cys Ser Gly Ser Gly Met Pro Ser Arg Gln
1 5 10 15

<210> 355

<211> 23

<212> PRT

<213> Homo sapiens

<400> 355

Tyr Ser Thr Cys Ser Leu Cys Gln Glu Glu Asn Glu Asp Val Val Arg
1 5 10 15

Asp Ala Leu Gln Gln Asn Pro
20

<210> 356

<211> 470

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (277)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (296)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (301)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

190

<221> SITE

<222> (306)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (324)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (431)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 356

Ser	Ala	Thr	Glu	His	Gly	Ala	Val	Cys	Cys	Ser	Cys	Arg	Arg	Val	Gly
1				5				10						15	

Arg	Arg	Gly	Glu	Pro	Pro	Gly	Ser	Ile	Lys	Gly	Leu	Val	Tyr	Ser	Ser
		20					25						30		

Asn	Phe	Gln	Asn	Val	Lys	Gln	Leu	Tyr	Ala	Leu	Val	Cys	Glu	Thr	Gln
	35					40						45			

Arg	Tyr	Ser	Ala	Val	Leu	Asp	Ala	Val	Ile	Ala	Ser	Ala	Gly	Leu	Leu
50					55						60				

Arg	Ala	Glu	Lys	Lys	Leu	Arg	Pro	His	Leu	Ala	Lys	Val	Leu	Val	Tyr
65				70					75						80

Glu	Leu	Leu	Leu	Gly	Lys	Gly	Phe	Arg	Gly	Gly	Gly	Gly	Arg	Trp	Lys
			85					90						95	

Ala	Leu	Leu	Gly	Arg	His	Gln	Ala	Arg	Leu	Lys	Ala	Glu	Leu	Ala	Arg
			100					105					110		

Leu	Lys	Val	His	Arg	Gly	Val	Ser	Arg	Asn	Glu	Asp	Leu	Leu	Glu	Val
	115						120					125			

Gly	Ser	Arg	Pro	Gly	Pro	Ala	Ser	Gln	Leu	Pro	Arg	Phe	Val	Arg	Val
130					135						140				

Asn	Thr	Leu	Lys	Thr	Cys	Ser	Asp	Asp	Val	Val	Asp	Tyr	Phe	Lys	Arg
145					150					155					160

Gln	Gly	Phe	Ser	Tyr	Gln	Gly	Arg	Ala	Ser	Ser	Leu	Asp	Asp	Leu	Arg
			165						170					175	

Ala	Leu	Lys	Gly	Lys	His	Phe	Leu	Leu	Asp	Pro	Leu	Met	Pro	Glu	Leu
			180					185						190	

Leu	Val	Phe	Pro	Ala	Gln	Thr	Asp	Leu	His	Glu	His	Pro	Leu	Tyr	Arg
	195						200						205		

Ala	Gly	His	Leu	Ile	Leu	Gln	Asp	Arg	Ala	Ser	Cys	Leu	Pro	Ala	Met
210						215						220			

Leu	Leu	Asp	Pro	Pro	Pro	Gly	Ser	His	Val	Ile	Asp	Ala	Cys	Ala	Ala
225						230					235				240

191

Pro	Gly	Asn	Lys	Thr	Ser	His	Leu	Ala	Ala	Leu	Lys	Asn	Gln	Gly	245	250	255	
Lys	Ile	Phe	Ala	Phe	Asp	Leu	Asp	Ala	Lys	Arg	Leu	Ala	Ser	Met	Ala	260	265	270
Thr	Leu	Leu	Ala	Xaa	Ala	Gly	Val	Ser	Cys	Cys	Glu	Leu	Ala	Glu	Glu	275	280	285
Asp	Phe	Leu	Ala	Val	Ser	Pro	Xaa	Asp	Pro	Arg	Tyr	Xaa	Glu	Val	His	290	295	300
Tyr	Xaa	Leu	Leu	Asp	Pro	Ser	Cys	Ser	Gly	Ser	Gly	Met	Pro	Ser	Arg	305	310	320
Gln	Leu	Glu	Xaa	Pro	Gly	Ala	Gly	Thr	Pro	Ser	Pro	Val	Arg	Leu	His	325	330	335
Ala	Leu	Ala	Gly	Phe	Gln	Gln	Arg	Ala	Leu	Cys	His	Ala	Leu	Thr	Phe	340	345	350
Pro	Ser	Leu	Gln	Arg	Leu	Val	Tyr	Ser	Thr	Cys	Ser	Leu	Cys	Gln	Glu	355	360	365
Glu	Asn	Glu	Asp	Val	Val	Arg	Asp	Ala	Leu	Gln	Gln	Asn	Pro	Gly	Ala	370	375	380
Phe	Arg	Leu	Ala	Pro	Ala	Leu	Pro	Ala	Trp	Pro	His	Arg	Gly	Leu	Ser	385	390	395
Thr	Phe	Pro	Gly	Ala	Glu	His	Cys	Leu	Arg	Ala	Ser	Pro	Glu	Thr	Thr	405	410	415
Leu	Ser	Ser	Gly	Phe	Phe	Val	Ala	Val	Ile	Glu	Arg	Val	Glu	Xaa	Pro	420	425	430
Ser	Ser	Ala	Ser	Gln	Ala	Lys	Ala	Ser	Ala	Pro	Glu	Arg	Thr	Pro	Ser	435	440	445
Pro	Ala	Pro	Lys	Arg	Lys	Lys	Arg	Gln	Gln	Arg	Ala	Ala	Ala	Gly	Ala	450	455	460
Cys	Thr	Pro	Pro	Cys	Thr											465	470	

<210> 357

<211> 429

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (236)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (255)

<223> Xaa equals any of the naturally occurring L-amino acids

<223> Xaa equals any of the naturally occurring L-amino acids

<223> Xaa equals any of the naturally occurring L-amino acids

<223> Xaa equals any of the naturally occurring L-amino acids

Tyr Glu Pro His Ser Thr His Ser Arg Glu Arg Ala Met Thr Ser His
1 5 10 15

Ala Arg Val Ser Leu Gly Pro Ser Arg Asp Pro Leu Glu Arg Pro His
20 25 30

Leu Ala Lys Val Leu Val Tyr Glu Leu Leu Leu Gly Lys Gly Phe Arg
35 40 45

Gly Gly Gly Gly Arg Trp Lys Ala Leu Leu Gly Arg His Gln Ala Arg
50 55 60

Leu Lys Ala Glu Leu Ala Arg Leu Lys Val His Arg Gly Val Ser Arg
65 70 75 80

Asn Glu Asp Leu Leu Glu Val Gly Ser Arg Pro Gly Pro Ala Ser Gln
85 90 95

Leu Pro Arg Phe Val Arg Val Asn Thr Leu Lys Thr Cys Ser Asp Asp
100 105 110

Val Val Asp Tyr Phe Lys Arg Gln Gly Phe Ser Tyr Gln Gly Arg Ala
115 120 125

Ser Ser Leu Asp Asp Leu Arg Ala Leu Lys Gly Lys His Phe Leu Leu
130 135 140

Asp Pro Leu Met Pro Glu Leu Leu Val Phe Pro Ala Gln Thr Asp Leu
145 150 155 160

His Glu His Pro Leu Tyr Arg Ala Gly His Leu Ile Leu Gln Asp Arg
165 170 175

Ala Ser Cys Leu Pro Ala Met Leu Leu Asp Pro Pro Pro Gly Ser His
180 185 190

Val	Ile	Asp	Ala	Cys	Ala	Ala	Pro	Gly	Asn	Lys	Thr	Ser	His	Leu	Ala
		195					200					205			

Ala Leu Leu Lys Asn Gln Gly Lys Ile Phe Ala Phe Asp Leu Asp Ala
210 215 220

193

Lys Arg Leu Ala Ser Met Ala Thr Leu Leu Ala Xaa Ala Gly Val Ser
 225 230 235 240
 Cys Cys Glu Leu Ala Glu Glu Asp Phe Leu Ala Val Ser Pro Xaa Asp
 245 250 255
 Pro Arg Tyr Xaa Glu Val His Tyr Xaa Leu Leu Asp Pro Ser Cys Ser
 260 265 270
 Gly Ser Gly Met Pro Ser Arg Gln Leu Glu Glu Pro Gly Ala Gly Thr
 275 280 285
 Pro Ser Pro Val Arg Leu His Ala Leu Ala Gly Phe Gln Gln Arg Ala
 290 295 300
 Leu Cys His Ala Leu Thr Phe Pro Ser Leu Gln Arg Leu Val Tyr Ser
 305 310 315 320
 Thr Cys Ser Leu Cys Gln Glu Glu Asn Glu Asp Val Val Arg Asp Ala
 325 330 335
 Leu Gln Gln Asn Pro Gly Ala Phe Arg Leu Ala Pro Ala Leu Pro Ala
 340 345 350
 Trp Pro His Arg Gly Leu Ser Thr Phe Pro Gly Ala Glu His Cys Leu
 355 360 365
 Arg Ala Ser Pro Glu Thr Thr Leu Ser Ser Gly Phe Phe Val Ala Val
 370 375 380
 Ile Glu Arg Val Glu Val Pro Ser Ser Ala Ser Gln Ala Lys Ala Ser
 385 390 395 400
 Ala Pro Glu Arg Thr Pro Ser Pro Ala Pro Lys Arg Lys Lys Arg Gln
 405 410 415
 Gln Xaa Ala Ala Ala Gly Ala Cys Thr Pro Pro Cys Thr
 420 425

<210> 358

<211> 245

<212> PRT

<213> Homo sapiens

<400> 358

Met Gly Thr His Ser Val Ser Gly Arg Phe Ser Lys Thr Ser Pro Pro
 1 5 10 15
 Tyr Cys Pro Pro Ser Ser Ser Leu Pro Gly Pro Ile Ser Ser Ile Gly
 20 25 30
 Phe Asn Lys Ser Leu His Glu Cys Leu Phe Ile Ser Glu Lys Glu Leu
 35 40 45
 Leu Pro Leu Pro Phe Pro Phe Pro Asp Leu Lys Ser Phe Ile Ser Tyr
 50 55 60
 Leu Thr Ser Met Leu Lys Pro Gly Pro Leu Ile Val Ser Leu Lys Ile

<210> 359

<211> 29

<212> PRT

<213> Homo sapiens

<400> 359

Cys Phe Ala Arg Leu Leu Lys Ser Pro Leu Phe Thr Pro Val Pro Thr
1 5 10 15

Glu Ile Ser Pro Phe Tyr Asn Cys Ala Tyr Tyr Ser Ala
20 25

<210> 360

<211> 111

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 360

Asn Arg Glu Gln Lys Ala Lys Ser Gln Leu Leu Arg Ser Gln Leu Tyr

195
 1 5 10 15
 Ser Thr Leu Asp Leu Pro Tyr Phe Phe Gln Cys Val Gly Thr Arg Cys
 20 25 30
 Thr Ala Val Cys Val Cys Val Cys Val Cys Val Cys Val Cys Xaa Tyr
 35 40 45
 Leu Pro Ile His Trp Gln Val Asn Leu His Leu Val Tyr Leu Ala Met
 50 55 60
 Leu Cys Phe Leu Pro Ile Pro Leu Leu Ser Ile Leu Ser Pro Gln Thr
 65 70 75 80
 Gln Ala Ser Arg Leu Leu Asp Glu Thr Val Arg Arg Lys His Phe Leu
 85 90 95
 Thr Tyr Pro Phe Gly Ile Ser Ser Ile Ile Thr Gln Ala Leu Leu
 100 105 110

<210> 361
 <211> 51
 <212> PRT
 <213> Homo sapiens

<400> 361
 Pro Gly Pro Glu Ala Gln Pro Trp Pro Gly Pro Asp Leu Pro Ala Val
 1 5 10 15
 Gly Ser Arg Gly Pro Gly Arg Leu Leu Ala Ala Val Ser Ala Pro Arg
 20 25 30
 Leu Gly Leu Gly Leu Ala Gly Ala Asp Pro Val Gly Pro Glu Ala Cys
 35 40 45

His Leu Pro
 50

<210> 362
 <211> 42
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (32)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 362
 Gly Arg Leu Arg Gly Pro Asp Glu Val Gly Ala Pro Phe His Pro Gly
 1 5 10 15
 Pro Ala Thr Pro Gly Leu Ala Asp Pro Leu Arg Pro Ala Glu Pro Xaa
 20 25 30

His Trp Leu Pro Ser Leu Trp Gly Pro Thr
 35 40

<210> 363

196

<211> 19
<212> PRT
<213> Homo sapiens

<400> 363
Pro Gly Pro Glu Ala Gln Pro Trp Pro Gly Pro Asp Leu Pro Ala Val
1 5 10 15

Gly Ser Arg

<210> 364
<211> 19
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (15)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 364
Ala Thr Pro Gly Leu Ala Asp Pro Leu Arg Pro Ala Glu Pro Xaa His
1 5 10 15

Trp Leu Pro

<210> 365
<211> 251
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (210)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (241)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 365
Gln Trp Pro Glu Lys Asp Pro Val Met Ala Ala Ser Ser Ile Ser Ser
1 5 10 15

Pro Trp Gly Lys His Val Phe Lys Ala Ile Leu Met Val Leu Val Ala
20 25 30

Leu Ile Leu Leu His Ser Ala Leu Ala Gln Ser Arg Arg Asp Phe Ala
35 40 45

Pro Pro Gly Gln Gln Lys Arg Glu Ala Pro Val Asp Val Leu Thr Gln
50 55 60

Ile Gly Arg Ser Val Arg Gly Thr Leu Asp Ala Trp Ile Gly Pro Glu
65 70 75 80

197
 Thr Met His Leu Val Ser Glu Ser Ser Ser Gln Val Leu Trp Ala Ile
 85 90 95
 Ser Ser Ala Ile Ser Val Ala Phe Phe Ala Leu Ser Gly Ile Ala Ala
 100 105 110
 Gln Leu Leu Asn Ala Leu Gly Leu Ala Gly Asp Tyr Leu Ala Gln Gly
 115 120 125
 Leu Lys Leu Ser Pro Gly Gln Val Gln Thr Phe Leu Leu Trp Gly Ala
 130 135 140
 Gly Ala Leu Val Val Tyr Trp Leu Leu Ser Leu Leu Leu Gly Leu Val
 145 150 155 160
 Leu Ala Leu Leu Gly Arg Ile Leu Trp Gly Leu Lys Leu Val Ile Phe
 165 170 175
 Leu Ala Gly Phe Val Ala Leu Met Arg Ser Val Pro Asp Pro Ser Thr
 180 185 190
 Arg Ala Leu Leu Leu Leu Ala Leu Leu Ile Leu Tyr Ala Leu Leu Ser
 195 200 205
 Arg Xaa Thr Gly Ser Arg Ala Ser Gly Ala Gln Leu Glu Ala Lys Val
 210 215 220
 Arg Gly Leu Glu Arg Gln Val Glu Glu Leu Arg Trp Arg Gln Arg Gln
 225 230 235 240
 Xaa Ala Lys Gly Ala Arg Ser Val Glu Glu Glu
 245 250

<210> 366

<211> 116

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (2)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (5)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (7)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (9)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 366

198

Glu Xaa Pro Arg Xaa Ile Xaa Gly Xaa Asn Ala Pro Gln Val Pro Val
 1 5 10 15

Arg Asn Ser Arg Val Asp Pro Arg Val Arg Pro Arg Val Arg Ser Leu
 20 25 30

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn Gly
 35 40 45

Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu Gly Leu
 50 55 60

Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser Ser Asn Lys
 65 70 75 80

Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu Asp Tyr Ile Ile
 85 90 95

Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val Ser Arg Asn Leu Gly
 100 105 110

Pro Lys Ile Ile
 115

<210> 367

<211> 12

<212> PRT

<213> Homo sapiens

<400> 367

Asn Ile Leu Leu Val Asn Leu Leu Val Ala Met Phe
 1 5 10

<210> 368

<211> 10

<212> PRT

<213> Homo sapiens

<400> 368

Gln Val Trp Lys Phe Gln Arg Tyr Phe Leu
 1 5 10

<210> 369

<211> 316

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (2)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (5)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (7)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (9)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (126)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (127)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (143)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (166)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (176)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (200)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (294)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (296)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (306)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <400> 369
 Glu Xaa Pro Arg Xaa Ile Xaa Gly Xaa Asn Ala Pro Gln Val Pro Val
 1 5 10 15
 Arg Asn Ser Arg Val Asp Pro Arg Val Arg Pro Arg Val Arg Ser Leu
 20 25 30

200

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn Gly
 35 40 45
 Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu Gly Leu
 50 55 60
 Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser Ser Asn Lys
 65 70 75 80
 Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu Asp Tyr Ile Ile
 85 90 95
 Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val Ser Arg Asn Leu Gly
 100 105 110
 Pro Lys Ile Ile Met Leu Gln Arg Met Leu Ile Asp Val Xaa Xaa Phe
 115 120 125
 Leu Phe Leu Phe Ala Val Trp Met Val Ala Phe Gly Val Ala Xaa Gln
 130 135 140
 Gly Ile Leu Arg Gln Asn Glu Gln Arg Trp Arg Trp Ile Phe Arg Ser
 145 150 155 160
 Val Ile Tyr Glu Pro Xaa Leu Ala Met Phe Gly Gln Val Pro Ser Xaa
 165 170 175
 Val Asp Gly Thr Thr Tyr Asp Phe Ala His Cys Thr Phe Thr Gly Asn
 180 185 190
 Glu Ser Lys Pro Leu Cys Val Xaa Leu Asp Glu His Asn Leu Pro Arg
 195 200 205
 Phe Pro Glu Trp Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser
 210 215 220
 Thr Asn Ile Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr
 225 230 235 240
 Val Gly Thr Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg
 245 250 255
 Tyr Phe Leu Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro
 260 265 270
 Phe Ile Val Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe Lys
 275 280 285
 Cys Cys Cys Lys Glu Xaa Asn Xaa Glu Ser Ser Val Cys Cys Ser Lys
 290 295 300
 Met Xaa Thr Met Arg Leu Trp His Gly Arg Val Ser
 305 310 315

<210> 370

<211> 129

<212> PRT

<213> Homo sapiens

201

<400> 370

Met Glu Phe Gln Asn Met Tyr Ile Gln Leu Phe Gly Phe Ser Phe Phe
1 5 10 15

Ile Val Ile Ile Val Arg Met Leu Leu Leu Gly Leu Cys Val Ser Ala
20 25 30

Arg Gln Pro Val Met Pro Arg Ala Thr Leu Trp Gly His Leu Ser Pro
35 40 45

Ala Trp Val Leu Val Pro Trp Thr Pro Arg Ala Cys Gly Gln Ala Ala
50 55 60

Pro Gly Arg Gly His Val Ala Ser Asp His Lys Ser Gly Leu Pro Trp
65 70 75 80

Pro Lys His Cys Ser Cys Leu His Pro Arg Ala Ser Gln Pro Cys Leu
85 90 95

Phe Ser Leu Asn Ser Asn Arg Thr Val Phe Thr Ala Ile Gln Arg Val
100 105 110

Ala Leu Gly Trp Thr Phe Trp Val Gln Ala Asn Leu Val Pro Arg Cys
115 120 125

Thr

<210> 371

<211> 417

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (54)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (90)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (109)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (111)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (121)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (135)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (137)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (139)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (188)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (205)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (223)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (249)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (252)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (322)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (348)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (402)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 371
Leu Leu Leu Cys Val Thr Gly Val Tyr Ser Tyr Gly Leu Met His Pro
1 5 10 15

203

Ile Pro Ser Ser Phe Met Ile Lys Ala Val Ser Ser Phe Leu Thr Ala
 20 25 30

Glu Glu Ala Ser Val Gly Asn Pro Glu Gly Ala Phe Met Lys Val Leu
 35 40 45

Gln Ala Arg Lys Asn Xaa Thr Ser Thr Glu Leu Ile Val Glu Pro Glu
 50 55 60

Glu Pro Ser Asp Ser Ser Gly Ile Asn Leu Ser Gly Phe Gly Ser Glu
 65 70 75 80

Gln Leu Asp Thr Asn Asp Glu Ser Asp Xaa Ile Ser Thr Leu Ser Tyr
 85 90 95

Ile Leu Pro Tyr Phe Ser Ala Val Asn Leu Asp Val Xaa Ser Xaa Leu
 100 105 110

Leu Pro Phe Ile Lys Leu Pro Thr Xaa Gly Asn Ser Leu Ala Lys Ile
 115 120 125

Gln Thr Val Gly Gln Asn Xaa Gln Xaa Val Xaa Arg Val Leu Met Gly
 130 135 140

Pro Arg Ser Ile Gln Lys Arg His Phe Lys Glu Val Gly Arg Gln Ser
 145 150 155 160

Ile Arg Arg Glu Gln Gly Ala Gln Ala Ser Val Glu Asn Ala Ala Glu
 165 170 175

Glu Lys Arg Leu Gly Ser Pro Ala Pro Arg Glu Xaa Glu Gln Pro His
 180 185 190

Thr Gln Gln Gly Pro Glu Lys Leu Ala Gly Asn Ala Xaa Tyr Thr Lys
 195 200 205

Pro Ser Phe Thr Gln Glu His Lys Ala Ala Val Ser Val Leu Xaa Pro
 210 215 220

Phe Ser Lys Gly Ala Pro Ser Thr Ser Ser Pro Ala Lys Ala Leu Pro
 225 230 235 240

Gln Val Arg Asp Arg Trp Lys Asp Xaa Thr His Xaa Ile Ser Ile Leu
 245 250 255

Glu Ser Ala Lys Ala Arg Val Thr Asn Met Lys Ala Ser Lys Pro Ile
 260 265 270

Ser His Ser Arg Lys Lys Tyr Arg Phe His Lys Thr Arg Ser Arg Met
 275 280 285

Thr His Arg Thr Pro Lys Val Lys Lys Ser Pro Lys Phe Arg Lys Lys
 290 295 300

Ser Tyr Leu Ser Arg Leu Met Leu Ala Asn Arg Pro Pro Phe Ser Ala
 305 310 315 320

Ala Xaa Ser Leu Ile Asn Ser Pro Ser Gln Gly Ala Phe Ser Ser Leu
 325 330 335

204

Gly Asp Leu Ser Pro Gln Glu Asn Pro Phe Leu Xaa Val Ser Ala Pro
340 345 350

Ser Glu His Phe Ile Glu Thr Thr Asn Ile Lys Asp Thr Thr Ala Arg
355 360 365

Asn Ala Leu Glu Glu Asn Val Phe Met Glu Asn Thr Asn Met Pro Glu
370 375 380

Val	Thr	Ile	Ser	Glu	Asn	Thr	Asn	Tyr	Asn	His	Pro	Pro	Glu	Ala	Asp
385					390					395					400

Ser Xaa Gly Thr Ala Phe Asn Leu Gly Pro Thr Val Lys Gln Thr Glu
405 410 415

Thr

```
<210> 372
<211> 94
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> SITE  
<222> (66)  
<223> Xaa equals any of the naturally occurring L-amino acids
```

<400> 372
Cys Phe Ser Asn Ala Pro Lys Val Ser Asp Glu Ala Val Lys Lys Asp
1 5 10 15

Ser Glu Leu Asp Lys His Leu Glu Ser Arg Val Glu Glu Ile Met Glu
20 25 30

Lys Ser Gly Glu Glu Gly Met Pro Asp Leu Ala His Val Met Arg Ile
35 40 45

Leu Ser Ala Glu Asn Ile Pro Asn Leu Pro Pro Gly Gly Gly Leu Ala
50 55 60

Gly Xaa Arg Asn Val Ile Glu Ala Val Tyr Ser Arg Leu Asn Pro His
65 70 75 80

Arg Glu Ser Asp Gly Gly Ala Gly Asp Leu Glu Asp Pro Trp
85 90

```
<210> 373
<211> 56
<212> PRT
<213> Homo sapiens
```

<400> 373
Cys Phe Ser Asn Ala Pro Lys Val Ser Asp Glu Ala Val Lys Lys Asp
1 5 10 15

Ser Glu Leu Asp Lys His Leu Glu Ser Arg Val Glu Glu Ile Met Glu
20 25 30

205

Lys Ser Gly Glu Glu Gly Met Pro Asp Leu Ala His Val Met Arg Ile
35 40 45

Leu Ser Ala Glu Asn Ile Pro Asn
50 55

<210> 374
<211> 26
<212> PRT
<213> Homo sapiens

<400> 374
Arg Asn Val Ile Glu Ala Val Tyr Ser Arg Leu Asn Pro His Arg Glu
1 5 10 15

Ser Asp Gly Gly Ala Gly Asp Leu Glu Asp
20 25

<210> 375
<211> 16
<212> PRT
<213> Homo sapiens

<400> 375
Asp Ser Glu Leu Asp Lys His Leu Glu Ser Arg Val Glu Glu Ile Met
1 5 10 15

<210> 376
<211> 24
<212> PRT
<213> Homo sapiens

<400> 376
Lys Ser Gly Glu Glu Gly Met Pro Asp Leu Ala His Val Met Arg Ile
1 5 10 15

Leu Ser Ala Glu Asn Ile Pro Asn
20

<210> 377
<211> 9
<212> PRT
<213> Homo sapiens

<400> 377
Cys Phe Ser Asn Ala Pro Lys Val Ser
1 5

<210> 378
<211> 69
<212> PRT
<213> Homo sapiens

<400> 378
Met Ser Arg Lys Ser Leu Ala Phe Pro Ile Ile Cys Ser Tyr Leu Cys

1 5 206 10 15
 Phe Leu Thr Val Ala Thr Cys Ser Ile Ala Cys Thr Thr Val Phe Phe
 20 25 30
 Ala Asn Leu Arg His Thr Arg Tyr Ile Cys Ile Glu Leu Ser Ala Leu
 35 40 45
 Glu Thr Ser Gly Val Ile Ser Pro Gln Ile Asn Asn Val Pro Glu Val
 50 55 60
 His Gly Lys Tyr Ser
 65

<210> 379
 <211> 16
 <212> PRT
 <213> Homo sapiens

<400> 379
 Ile Gln Lys Met Thr Arg Val Arg Val Val Asp Asn Ser Ala Leu Gly
 1 5 10 15

<210> 380
 <211> 14
 <212> PRT
 <213> Homo sapiens

<400> 380
 Pro Arg Cys Ile His Val Tyr Lys Lys Asn Gly Val Gly Lys
 1 5 10

<210> 381
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 381
 Gly Asp Gln Ile Leu Leu Ala Ile Lys Gly Gln Lys Lys Lys Ala
 1 5 10 15

<210> 382
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 382
 Asn Pro Val Gly Thr Arg Ile Lys Thr Pro Ile Pro Thr Ser Leu
 1 5 10 15

<210> 383
 <211> 171
 <212> PRT
 <213> Homo sapiens

<220>

207

<221> SITE

<222> (20)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 383

Val Leu Ile Pro Ser Phe Ser Ser Ser Phe Leu Cys Ser Arg Gly Gly
 1 5 10 15

Pro Leu Pro Xaa Asp Leu Ser Trp Asp Pro Met Ala Phe Phe Thr Gly
 20 25 30

Leu Trp Gly Pro Phe Thr Cys Val Ser Arg Val Leu Ser His His Cys
 35 40 45

Phe Ser Thr Thr Gly Ser Leu Ser Ala Ile Gln Lys Met Thr Arg Val
 50 55 60

Arg Val Val Asp Asn Ser Ala Leu Gly Asn Ser Pro Tyr His Arg Ala
 65 70 75 80

Pro Arg Cys Ile His Val Tyr Lys Lys Asn Gly Val Gly Lys Val Gly
 85 90 95

Asp Gln Ile Leu Leu Ala Ile Lys Gly Gln Lys Lys Lys Ala Leu Ile
 100 105 110

Val Gly His Cys Met Pro Gly Pro Arg Met Thr Pro Arg Phe Asp Ser
 115 120 125

Asn Asn Val Val Leu Ile Glu Asp Asn Gly Asn Pro Val Gly Thr Arg
 130 135 140

Ile Lys Thr Pro Ile Pro Thr Ser Leu Arg Lys Arg Glu Gly Glu Tyr
 145 150 155 160

Ser Lys Val Leu Ala Ile Ala Gln Asn Phe Val
 165 170

<210> 384

<211> 171

<212> PRT

<213> Homo sapiens

<400> 384

Ala Arg Val Val Gln Pro Ala Ala Arg Ala Gly Met Trp Ala Gly Gly
 1 5 10 15

Arg Ser Ser Cys Gln Ala Glu Val Leu Arg Ala Thr Arg Gly Gly Ala
 20 25 30

Ala Arg Gly Asn Ala Ala Pro Gly Arg Ala Leu Glu Met Val Pro Gly
 35 40 45

Ala Ala Gly Trp Cys Cys Leu Val Leu Trp Leu Pro Ala Cys Val Ala
 50 55 60

Ala His Gly Phe Arg Ile His Asp Tyr Leu Tyr Phe Gln Val Leu Ser
 65 70 75 80

208
 Pro Gly Asp Ile Arg Tyr Ile Phe Thr Ala Thr Pro Ala Lys Asp Phe
 85 90 95
 Gly Gly Ile Phe His Thr Arg Tyr Glu Gln Ile His Leu Val Pro Ala
 100 105 110
 Glu Pro Pro Glu Ala Cys Gly Glu Leu Ser Asn Gly Phe Phe Ile Gln
 115 120 125
 Asp Gln Ile Ala Leu Val Glu Arg Gly Gly Cys Ser Phe Leu Ser Lys
 130 135 140
 Thr Arg Val Val Gln Glu His Gly Gly Arg Ala Val Ile Ile Ser Asp
 145 150 155 160
 Asn Ala Leu Thr Met Thr Ala Ser Thr Trp Arg
 165 170

 <210> 385
 <211> 187
 <212> PRT
 <213> Homo sapiens

 <400> 385
 Ile Ala Thr Ala Ala Leu Phe Phe Phe Phe Tyr Cys Gln Val Ala Gly
 1 5 10 15
 Phe Ile Gly Lys Gly Gln Ser Leu Arg Ser Trp Val Pro Gln Arg Leu
 20 25 30
 Leu Gly Leu Glu Pro Gln Leu Gln Pro Met Gln Gln Ser Arg Leu Leu
 35 40 45
 Leu Pro Phe Leu Phe Phe Leu Leu Glu Gly Cys Ala Pro Ser Ser Leu
 50 55 60
 Gly Pro Gly Ala Ala Pro Gly Ser Gly His Ser Leu Gly Pro Pro Gly
 65 70 75 80
 Ser Pro Gly Ala Pro Gly Pro Gln Pro Ala Val Gly Pro Ser Ser Pro
 85 90 95
 Cys Gln Pro Gly Pro Ser Pro Ser Ser Pro Ala Ala Ala Ala Ala Ser
 100 105 110
 Ser Gln Ser Ser Val Ala Ser Trp Pro Cys Thr Leu Arg Cys Ala Ala
 115 120 125
 Pro Ser Pro Asp Ala Ser Ala Leu Arg Pro Ala Ala Ser Pro Ala Ala
 130 135 140
 Thr Pro Ala Trp Ser Pro Gly Ser Gly Thr Ile Arg Val Leu Arg Pro
 145 150 155 160
 Pro Ala Pro Ala Ala Ala Pro Ala Thr Ala Ile Thr Asn Arg Gly Pro
 165 170 175
 Pro Arg Arg Arg Arg Arg Asn Ala Arg Thr Ala
 180 185

<210> 386
 <211> 194
 <212> PRT
 <213> Homo sapiens

<400> 386
 Glu Arg Pro Pro Arg Arg Thr Gly Thr Pro Val Ala Arg Pro Arg
 1 5 10 15
 Gly Pro Pro Asp Pro Ala Val Ala Ala Gly Thr Ala Leu Arg Ala Lys
 20 25 30
 Gln Phe Ala Arg Tyr Gly Ala Ala Ser Gly Val Val Pro Gly Ser Leu
 35 40 45
 Trp Pro Ser Pro Glu Gln Leu Arg Glu Leu Glu Ala Glu Glu Arg Glu
 50 55 60
 Trp Tyr Pro Ser Leu Ala Thr Met Gln Glu Ser Leu Arg Val Lys Gln
 65 70 75 80
 Leu Ala Glu Glu Gln Lys Arg Arg Glu Arg Glu Gln His Ile Ala Glu
 85 90 95
 Cys Met Ala Lys Met Pro Gln Met Ile Val Asn Trp Gln Gln Gln Gln
 100 105 110
 Arg Glu Asn Trp Glu Lys Ala Gln Ala Asp Lys Glu Arg Arg Ala Arg
 115 120 125
 Leu Gln Ala Glu Ala Gln Glu Leu Leu Gly Tyr Gln Val Asp Pro Arg
 130 135 140
 Ser Ala Arg Phe Gln Glu Leu Leu Gln Asp Leu Glu Lys Lys Glu Arg
 145 150 155 160
 Asn Pro Gln Gly Gly Lys Thr Glu Thr Glu Glu Gly Gly Ala Thr Ala
 165 170 175
 Ala Leu Ala Ala Ala Val Ala Gln Asp Pro Ala Ala Ser Gly Ala Pro
 180 185 190
 Ser Ser

<210> 387
 <211> 113
 <212> PRT
 <213> Homo sapiens

<400> 387
 Tyr Gln Ser Leu Ala Glu Thr Gln Gln Lys Lys Glu Asn Phe Arg Pro
 1 5 10 15
 Ile Ser Leu Lys Asn Thr Asp Ala Lys Ile Leu Asn Lys Ile Leu Ala
 20 25 30
 Asn Gln Ile Gln Gln His Ile Lys Lys Leu Ile His Asn Asp Arg Val

210

35 40 45

Gly Phe Ile Pro Glu Met Gln Gly Trp Phe Asn Ile Cys Lys Ser Ile
 50 55 60

Asn Ile Val His His Ile Asn Arg Thr Lys Asp Lys Asn His Met Ile
 65 70 75 80

Ile Ser Ile Asp Ala Glu Lys Ala Phe Asp Lys Ile Arg Gln Ser Phe
 85 90 95

Met Leu Lys Thr Leu Asn Lys Leu Gly Ile His Gly Met Tyr Leu Gly
 100 105 110

Arg

<210> 388
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 388
 Lys Lys Glu Asn Phe Arg Pro Ile Ser Leu Lys Asn Thr Asp Ala Lys
 1 5 10 15

Ile Leu Asn Lys Ile Leu Ala Asn Gln Ile Gln Gln His Ile Lys Lys
 20 25 30

Leu Ile His Asn Asp Arg Val Gly Phe Ile Pro Glu Met Gln Gly Trp
 35 40 45

Phe Asn Ile Cys Lys Ser Ile Asn Ile Val His His Ile Asn Arg Thr
 50 55 60

Lys Asp Lys Asn His Met Ile Ile Ser Ile Asp Ala Glu Lys Ala Phe
 65 70 75 80

Asp Lys Ile Arg Gln Ser Phe Met Leu Lys Thr Leu Asn Lys Leu Gly
 85 90 95

Ile His Gly Met Tyr
 100

<210> 389
 <211> 11
 <212> PRT
 <213> Homo sapiens

<400> 389
 Asp Ala Lys Ile Leu Asn Lys Ile Leu Ala Asn
 1 5 10

<210> 390
 <211> 10
 <212> PRT
 <213> Homo sapiens

<400> 390

211

Ile Gln Gln His Ile Lys Lys Leu Ile His
 1 5 10

<210> 391

<211> 19

<212> PRT

<213> Homo sapiens

<400> 391

Lys Asp Lys Asn His Met Ile Ile Ser Ile Asp Ala Glu Lys Ala Phe
 1 5 10 15

Asp Lys Ile

<210> 392

<211> 10

<212> PRT

<213> Homo sapiens

<400> 392

Met Leu Lys Thr Leu Asn Lys Leu Gly Ile
 1 5 10

<210> 393

<211> 10

<212> PRT

<213> Homo sapiens

<400> 393

Lys Lys Glu Asn Phe Arg Pro Ile Ser Leu
 1 5 10

<210> 394

<211> 85

<212> PRT

<213> Homo sapiens

<400> 394

Trp Thr Met Phe Ile Asp Leu His Met Leu Asn Gln Pro Cys Ile Ser
 1 5 10 15

Gly Met Lys Pro Thr Arg Ser Leu Trp Ile Ser Phe Leu Met Cys Cys
 20 25 30

Trp Ile Trp Phe Ala Asn Ile Leu Leu Arg Ile Phe Ala Ser Val Phe
 35 40 45

Phe Arg Asp Ile Gly Leu Lys Phe Ser Phe Phe Cys Cys Val Ser Ala
 50 55 60

Arg Leu Trp Tyr Gln Asp Asp Ala Gly Leu Ile Asn Glu Leu Gly Arg
 65 70 75 80

Ile Pro Ser Phe Tyr
 85

<210> 395

<211> 72

212

<212> PRT

<213> Homo sapiens

<400> 395

Glu Arg Pro Glu Glu Gly Thr Glu Pro Ser Pro Ser Pro Val Ala Glu
1 5 10 15

Gln Ala Ser Val Ser Met Thr Pro Val Phe Arg Ala Trp Gly Leu Trp
20 25 30

Val Tyr Val Leu Pro Thr Gly Phe Pro Gly Pro Cys Cys Met Met Leu
35 40 45

Leu Glu Leu Phe Pro Lys Glu Ser Val Pro Gln Ala Tyr Gln Gly Ile
50 55 60

Leu Leu Tyr Leu His Phe Gly Phe
65 70

<210> 396

<211> 123

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (23)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (27)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (32)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (106)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 396

Arg Gly Glu Val Pro His Gln Pro His Pro Thr Arg Arg Thr Val Val
1 5 10 15

Ser Gly Gln Ala Pro Trp Xaa Pro Gly Pro Xaa Ala Leu Gly Gln Xaa
20 25 30

Val Glu Thr Ala Ala Gly Met Gly Met Pro Leu Val Thr Val Thr Ala
35 40 45

Ala Thr Phe Pro Thr Leu Ser Cys Pro Pro Arg Ala Trp Pro Glu Val
50 55 60

Glu Ala Pro Glu Ala Pro Ala Leu Pro Val Val Pro Glu Leu Pro Glu
65 70 75 80

213

Val Pro Met Glu Met Pro Leu Val Leu Pro Pro Glu Leu Glu Leu Leu
85 90 95

Ser Leu Glu Ala Val His Arg Tyr Gln Xaa Gly Gly Thr Leu Met Gly
100 105 110

Trp Thr Arg Ala Glu Ala Ser Ala Asn Gly Ser
115 120

<210> 397

<211> 133

<212> PRT

<213> Homo sapiens

<400> 397

Met Val Leu Asp Pro Tyr Arg Ala Val Ala Leu Glu Leu Gln Ala Asn
1 5 10 15

Arg Glu Pro Asp Phe Ser Ser Leu Val Ser Pro Leu Ser Pro Arg Arg
20 25 30

Met Ala Ala Arg Val Phe Tyr Leu Leu Leu Gly Glu Cys Met His Val
35 40 45

Cys Val Cys Met Trp Gly Arg Asp Thr Glu Thr Arg Gly Pro Tyr Arg
50 55 60

Asp Ser Pro Asp Leu Pro Ser Pro Arg Leu Leu Thr Ser Ala Leu Ser
65 70 75 80

Ala Thr Asp Ser Ser Arg Glu Thr Arg Lys Ala Ile Trp Ser Pro Pro
85 90 95

Asp Pro Ala Gly Ala Gln Ile Pro Leu Arg Leu Glu Ser Ile Tyr Lys
100 105 110

Ala Ala Arg Lys Pro Ala Thr Ser Ser Lys Pro Arg Arg Ala Ser Leu
115 120 125

Lys Lys Lys Lys Lys
130

<210> 398

<211> 11

<212> PRT

<213> Homo sapiens

<400> 398

Ala Phe Arg Asn Leu Pro Asn Leu Arg Ile Leu
1 5 10

<210> 399

<211> 13

<212> PRT

<213> Homo sapiens

<400> 399

Ala Phe Gln Gly Leu Phe His Leu Phe Glu Leu Arg Leu

1 5 214
 10
 <210> 400
 <211> 206
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (3)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <400> 400
 Asn Lys Xaa Ile Leu Glu Val Pro Ser Ala Arg Thr Thr Arg Ile Met
 1 5 10 15
 Gly Asp His Leu Asp Leu Leu Leu Gly Val Val Leu Met Ala Gly Pro
 20 25 30
 Val Phe Gly Ile Pro Ser Cys Ser Phe Asp Gly Arg Ile Ala Phe Tyr
 35 40 45
 Arg Phe Cys Asn Leu Thr Gln Val Pro Gln Val Leu Asn Thr Thr Glu
 50 55 60
 Arg Leu Leu Leu Ser Phe Asn Tyr Ile Arg Thr Val Thr Ala Ser Ser
 65 70 75 80
 Phe Pro Phe Leu Glu Gln Leu Gln Leu Leu Glu Leu Gly Ser Gln Tyr
 85 90 95
 Thr Pro Leu Thr Ile Asp Lys Glu Ala Phe Arg Asn Leu Pro Asn Leu
 100 105 110
 Arg Ile Leu Asp Leu Gly Ser Ser Lys Ile Tyr Phe Leu His Pro Asp
 115 120 125
 Ala Phe Gln Gly Leu Phe His Leu Phe Glu Leu Arg Leu Tyr Phe Cys
 130 135 140
 Gly Leu Ser Asp Ala Val Leu Lys Asp Gly Tyr Phe Arg Asn Leu Lys
 145 150 155 160
 Ala Leu Thr Arg Leu Asp Leu Ser Lys Asn Gln Ile Arg Ser Leu Tyr
 165 170 175
 Leu His Pro Ser Phe Gly Lys Leu Asn Ser Leu Lys Ser Ile Asp Phe
 180 185 190
 Ser Ser Asn Gln Ile Phe Leu Val Cys Glu His Glu Leu Glu
 195 200 205

 <210> 401
 <211> 261
 <212> PRT
 <213> Homo sapiens

 <400> 401
 Ala His Ala Ala Leu Gln Leu Ser Leu Arg Thr Cys Gly Pro Cys Ser

1	5	10	15
Ser Pro Tyr Pro His Ala Gly Leu Ala Ala Leu Leu Thr His Met Trp	20	25	30
Ala Leu Gln Leu Ser Leu Pro Thr Cys Gly Leu Ala Ala Leu Leu Thr	35	40	45
His Met Arg Pro Cys Ser Ser Pro Tyr Pro His Ala Gly Leu Ala Ala	50	55	60
Leu Leu Thr His Met Gly Pro Cys Arg Ser Pro Tyr Pro His Gly Gly	65	70	75
Leu Ala Ala Val Leu Thr His Met Arg Ala Leu Gln Leu Ser Leu Pro	85	90	95
Thr Trp Gly Leu Ala Ala Leu Leu Thr His Met Arg Pro Cys Ser Ser	100	105	110
Pro Tyr Pro His Ala Gly Leu Ala Cys Cys Trp Leu Trp Ser Leu Ser	115	120	125
Ser His Arg Ser Leu Gln Val Gln Ala Thr His Arg Leu Val Val Arg	130	135	140
Thr Ile Lys Asp Arg Val Met Leu Lys Val Leu Pro Gln Thr Arg Arg	145	150	155
Arg Gly Pro Phe Leu Ser Ser Cys Arg Asn Asp Val Met Arg Asn Cys	165	170	175
Val Pro Arg His Ala Val Leu Val Thr Thr Cys Val Phe Val Ser Phe	180	185	190
Pro Thr His Cys Lys Val Gly Ile Thr Gly Pro Ile Thr Gln Val Lys	195	200	205
Gln Lys Pro Gly Asn His Ser Ser Pro Cys Pro Val Ile Gln Leu Val	210	215	220
Ala Lys Ala Glu Phe Glu Leu Met Leu Pro Ser Val Pro Lys Pro Val	225	230	235
Tyr Leu Thr Leu Val Leu Ser Cys Trp Cys Leu Cys Asp Val Pro Cys	245	250	255
Leu Ser Val Ser Leu	260		

<210> 402

<211> 17

<212> PRT

<213> Homo sapiens

<400> 402

Leu Ala Cys Cys Trp Leu Trp Ser Leu Ser Ser His Arg Ser Leu Gln
1 5 10 15

216

Val

<210> 403

<211> 67

<212> PRT

<213> Homo sapiens

<400> 403

Met Gly Glu Ala Ser Pro Pro Ala Pro Ala Arg Arg His Leu Leu Val
 1 5 10 15

Leu Leu Leu Leu Leu Ser Thr Leu Val Ile Pro Ser Ala Ala Ala Pro
 20 25 30

Ile His Asp Ala Asp Ala Gln Glu Ser Ser Leu Gly Leu Thr Gly Leu
 35 40 45

Gln Ser Leu Leu Gln Gly Phe Ser Arg Leu Phe Leu Lys Val Thr Cys
 50 55 60

Phe Gly Ala
 65

<210> 404

<211> 90

<212> PRT

<213> Homo sapiens

<400> 404

Met Leu Val Val Ser Thr Val Ile Ile Val Phe Trp Glu Phe Ile Asn
 1 5 10 15

Ser Thr Glu Gly Ser Phe Leu Trp Ile Tyr His Ser Lys Asn Pro Glu
 20 25 30

Val Asp Asp Ser Ser Ala Gln Lys Gly Trp Trp Phe Leu Ser Trp Phe
 35 40 45

Asn Asn Gly Ile His Asn Tyr Gln Gln Gly Glu Glu Asp Ile Asp Lys
 50 55 60

Glu Lys Gly Arg Glu Glu Thr Lys Gly Arg Lys Met Thr Gln Gln Ser
 65 70 75 80

Phe Gly Tyr Gly Thr Gly Leu Ile Gln Thr
 85 90

<210> 405

<211> 18

<212> PRT

<213> Homo sapiens

<400> 405

Phe Pro Gly Arg Thr His Ala Ser Gly Asn Val Lys Gly Lys Val Ile
 1 5 10 15

Leu Ser

217

<210> 406

<211> 106

<212> PRT

<213> Homo sapiens

<400> 406

Ala Asp Gln Glu Lys Ile Arg Asn Val Lys Gly Lys Val Ile Leu Ser
 1 5 10 15

Met Leu Val Val Ser Thr Val Ile Ile Val Phe Trp Glu Phe Ile Asn
 20 25 30

Ser Thr Glu Gly Ser Phe Leu Trp Ile Tyr His Ser Lys Asn Pro Glu
 35 40 45

Val Asp Asp Ser Ser Ala Gln Lys Gly Trp Trp Phe Leu Ser Trp Phe
 50 55 60

Asn Asn Gly Ile His Asn Tyr Gln Gln Gly Glu Glu Asp Ile Asp Lys
 65 70 75 80

Glu Lys Gly Arg Glu Glu Thr Lys Gly Arg Lys Met Thr Gln Gln Ser
 85 90 95

Phe Gly Tyr Gly Thr Gly Leu Ile Gln Thr
 100 105

<210> 407

<211> 236

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (50)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 407

Met Gln Ser Pro Leu Val Glu Cys Pro Pro Pro Ser Ile His Tyr Trp
 1 5 10 15

Pro Ser Val Pro Ala Gly Ala Gln Gly Ala Cys Ser Pro Met Phe His
 20 25 30

Ala Ala Gly Trp Ser Arg Ser Gln Pro Asn Gly Glu Ile Pro Ala Ser
 35 40 45

Ser Xaa Gly His Leu Ser Ile Gln Arg Ala Ala Leu Val Val Leu Glu
 50 55 60

Asn Tyr Tyr Lys Asp Phe Thr Ile Tyr Asn Pro Asn Leu Leu Thr Ala
 65 70 75 80

Ser Lys Phe Arg Ala Ala Lys His Met Ala Gly Leu Lys Val Tyr Asn
 85 90 95

Val Asp Gly Pro Ser Asn Asn Ala Thr Gly Gln Ser Arg Ala Met Ile
 100 105 110

218

Ala Ala Ala Ala Arg Arg Arg Asp Ser Ser His Asn Glu Leu Tyr Tyr
 115 120 125

Glu Glu Ala Glu His Glu Arg Arg Val Lys Lys Arg Lys Ala Arg Leu
 130 135 140

Val Val Ala Val Glu Glu Ala Phe Ile His Ile Gln Arg Leu Gln Ala
 145 150 155 160

Glu Glu Gln Gln Lys Ala Pro Gly Glu Val Met Asp Pro Arg Glu Ala
 165 170 175

Ala Gln Ala Ile Phe Pro Ser Met Ala Arg Ala Leu Gln Lys Tyr Leu
 180 185 190

Arg Ile Thr Arg Gln Gln Asn Tyr His Ser Met Glu Ser Ile Leu Gln
 195 200 205

Ala Pro Gly Leu Leu His His Gln Arg His Asp Pro Gln Gly Leu Pro
 210 215 220

Arg Thr Val Pro Gln Cys Gly Pro His Pro Ala Ile
 225 230 235

<210> 408

<211> 23

<212> PRT

<213> Homo sapiens

<400> 408

Leu Ser Ile Gln Arg Ala Ala Leu Val Val Leu Glu Asn Tyr Tyr Lys
 1 5 10 15

Asp Phe Thr Ile Tyr Asn Pro
 20

<210> 409

<211> 15

<212> PRT

<213> Homo sapiens

<400> 409

Asp Ser Ser His Asn Glu Leu Tyr Tyr Glu Glu Ala Glu His Glu
 1 5 10 15

<210> 410

<211> 18

<212> PRT

<213> Homo sapiens

<400> 410

Phe Pro Ser Met Ala Arg Ala Leu Gln Lys Tyr Leu Arg Ile Thr Arg
 1 5 10 15

Gln Gln

<210> 411

219

<211> 140

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (117)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 411

Met Ala Phe Lys Leu Leu Ile Leu Leu Ile Gly Thr Trp Ala Leu Phe
 1 5 10 15

Phe Arg Lys Arg Arg Ala Asp Met Pro Arg Val Phe Val Phe Arg Ala
 20 25 30

Leu Leu Leu Val Leu Ile Phe Leu Phe Cys Gly Phe Pro Ile Gly Phe
 35 40 45

Phe Thr Gly Ser Ala Phe Trp Thr Leu Gly Asn Arg Asn Tyr Gln Gly
 50 55 60

Ile Val Gln Tyr Ala Val Ser Pro Cys Gly Met Pro Ser Ser Phe His
 65 70 75 80

Pro Leu Leu Ala Ile Arg Pro Cys Trp Ser Ser Gly Ser Leu Gln Pro
 85 90 95

Asn Val Pro Arg Cys Arg Leu Val Pro Leu Pro Thr Glu Trp Gly Asn
 100 105 110

Pro Arg Phe Gln Xaa Gly Thr Pro Glu Tyr Pro Ala Ser Ser Ile Gly
 115 120 125

Gly Pro Arg Lys Leu Leu Gln Arg Phe His His Leu
 130 135 140

<210> 412

<211> 37

<212> PRT

<213> Homo sapiens

<400> 412

Met Gly Leu Pro Val Ser Trp Ala Pro Pro Ala Leu Trp Val Leu Gly
 1 5 10 15

Cys Cys Ala Leu Leu Leu Ser Leu Trp Ala Leu Cys Thr Ala Cys Arg
 20 25 30

Ser Pro Arg Thr Leu
 35

<210> 413

<211> 20

<212> PRT

<213> Homo sapiens

<400> 413

Ile Tyr Gly Lys Thr Gly Gln Pro Asp Lys Ile Tyr Val Glu Leu His

1 5 220 15
 10
 Gln Asn Ser Pro
 20
 <210> 414
 <211> 16
 <212> PRT
 <213> Homo sapiens
 <400> 414
 Phe Leu Glu Pro Leu Ser Gly Leu Tyr Thr Cys Thr Leu Ser Tyr Lys
 1 5 10 15
 <210> 415
 <211> 16
 <212> PRT
 <213> Homo sapiens
 <400> 415
 Leu Gln Val Val Arg Leu Asp Ser Cys Arg Pro Gly Phe Gly Lys Asn
 1 5 10 15
 <210> 416
 <211> 12
 <212> PRT
 <213> Homo sapiens
 <400> 416
 Cys Val Ser Val Leu Thr Tyr Gly Ala Lys Ser Cys
 1 5 10
 <210> 417
 <211> 308
 <212> PRT
 <213> Homo sapiens
 <400> 417
 Pro Ala Lys Gly Glu Gly Cys Arg Arg Leu His Asp His Pro His Ile
 1 5 10 15
 Trp Arg Leu Leu Trp Ala His Ser Asp Pro Asp Pro Leu Pro Thr Gln
 20 25 30
 Pro Arg Ala Glu Gln Gly Glu Thr Glu Phe Cys Val Pro Val Gly Pro
 35 40 45
 Leu Cys His Asp Trp His Pro Leu Pro Val Asp Val Leu Ala Gln Leu
 50 55 60
 Gln Leu Ser His Ile Leu Pro Trp Gly Gln Pro Ala Pro Ser Arg His
 65 70 75 80

```

<210> 418
<211> 108
<212> PRT
<213> Homo sapiens

<400> 418
Pro Arg Val Arg Thr Arg Ala Pro Val Val Pro Pro Ala Gly His Arg
 1             5             10             15

Ala Leu Ser Pro Ala Gly Val Leu Leu Ala Val Pro Ala Met Leu Ser
      20             25             30

Leu Asp Phe Leu Asp Asp Val Arg Arg Met Asn Lys Arg Gln Val Ser
 35             40             45

```

222

Leu Ser Val Leu Phe Phe Ser Trp Leu Phe Leu Ser Leu Arg Gly Cys
50 55 60

Cys Cys Gly Ala Arg Arg Thr Pro Gly Phe Trp Cys Glu Gly Leu Ser
65 70 75 80

Trp Ser Asp Thr Arg Val Ile Arg Phe Leu Trp Arg Leu Trp Pro Glu
85 90 95

Ala Ala Leu Ser Ala Ser Leu Phe Leu Thr Pro Asn
100 105

<210> 419

<211> 16

<212> PRT

<213> Homo sapiens

<400> 419

His Ala Ser Ala Trp Asn Leu Ile Leu Leu Thr Val Phe Thr Leu Ser
1 5 10 15

<210> 420

<211> 24

<212> PRT

<213> Homo sapiens

<400> 420

Val Tyr Ala Ala Leu Gly Ala Gly Val Phe Thr Leu Phe Leu Ala Leu
1 5 10 15

Asp Thr Gln Leu Leu Met Gly Asn
20

<210> 421

<211> 18

<212> PRT

<213> Homo sapiens

<400> 421

Glu Glu Tyr Ile Phe Gly Ala Leu Asn Ile Tyr Leu Asp Ile Ile Tyr
1 5 10 15

Ile Phe

<210> 422

<211> 26

<212> PRT

<213> Homo sapiens

<400> 422

Trp Asn Leu Ile Leu Leu Thr Val Phe Thr Leu Ser Met Ala Tyr Leu
1 5 10 15

Thr Gly Met Leu Ser Ser Tyr Tyr Asn Thr

20 223
25

<210> 423
 <211> 11
 <212> PRT
 <213> Homo sapiens

<400> 423
 Thr Leu Ser Leu Leu Val Ser Leu His Thr Val
 1 5 10

<210> 424
 <211> 241
 <212> PRT
 <213> Homo sapiens

<400> 424
 Met Ser Ser Ser Gly Thr Ser Asp Ala Ser Pro Ser Gly Ser Pro Val
 1 5 10 15

Leu Ala Ser Tyr Lys Pro Ala Pro Pro Lys Asp Lys Leu Pro Glu Thr
 20 25 30

Pro Arg Arg Arg Met Lys Lys Ser Leu Ser Ala Pro Leu His Pro Glu
 35 40 45

Phe Glu Glu Val Tyr Arg Phe Gly Ala Glu Ser Arg Lys Leu Leu Leu
 50 55 60

Arg Glu Pro Val Asp Ala Met Pro Asp Pro Thr Pro Phe Leu Leu Ala
 65 70 75 80

Arg Glu Ser Ala Glu Val His Leu Ile Lys Glu Arg Pro Leu Val Ile
 85 90 95

Pro Pro Ile Ala Ser Asp Arg Ser Gly Glu Gln His Ser Pro Ala Arg
 100 105 110

Glu Lys Pro His Lys Ala His Val Gly Val Ala His Arg Ile His His
 115 120 125

Ala Thr Pro Pro Gln Pro Ala Arg Gly Glu Asp Pro Gly Gly Arg Pro
 130 135 140

Gly Glu Arg Arg Gln Gly Gly Glu Glu Ala Leu Arg Asp Gly Gln Asn
 145 150 155 160

Cys Val Lys Pro Ala Val Pro His Pro Ala Leu Ser Met His Cys Glu
 165 170 175

His His Trp Glu Ile Ser Ala Thr Pro Phe Leu Phe Asn Pro Met His
 180 185 190

Ala Lys His Phe Ser His Leu Pro Thr His Ser Pro Ser Ala Ser Leu
 195 200 205

Ala Leu Phe Phe Thr Pro Lys Tyr Asp Arg Val Pro Ala Ala Glu Tyr
 210 215 220

224

Val Phe Pro Asn Cys Cys Gly Gln Thr Pro Val Cys Arg Ile Ala Cys
 225 230 235 240

Phe

<210> 425

<211> 85

<212> PRT

<213> Homo sapiens

<400> 425

Met Ser Ser Ser Gly Thr Ser Asp Ala Ser Pro Ser Gly Ser Pro Val
 1 5 10 15

Leu Ala Ser Tyr Lys Pro Ala Pro Pro Lys Asp Lys Leu Pro Glu Thr
 20 25 30

Pro Arg Arg Arg Met Lys Lys Ser Leu Ser Ala Pro Leu His Pro Glu
 35 40 45

Phe Glu Glu Val Tyr Arg Phe Gly Ala Glu Ser Arg Lys Leu Leu Leu
 50 55 60

Arg Glu Pro Val Asp Ala Met Pro Asp Pro Thr Pro Phe Leu Leu Ala
 65 70 75 80

Arg Glu Ser Ala Glu
 85

<210> 426

<211> 63

<212> PRT

<213> Homo sapiens

<400> 426

Val His Leu Ile Lys Glu Arg Pro Leu Val Ile Pro Pro Ile Ala Ser
 1 5 10 15

Asp Arg Ser Gly Glu Gln His Ser Pro Ala Arg Glu Lys Pro His Lys
 20 25 30

Ala His Val Gly Val Ala His Arg Ile His His Ala Thr Pro Pro Gln
 35 40 45

Pro Ala Arg Gly Glu Asp Pro Gly Gly Arg Pro Gly Glu Arg Arg
 50 55 60

<210> 427

<211> 93

<212> PRT

<213> Homo sapiens

<400> 427

Gln Gly Gly Glu Glu Ala Leu Arg Asp Gly Gln Asn Cys Val Lys Pro
 1 5 10 15

Ala Val Pro His Pro Ala Leu Ser Met His Cys Glu His His Trp Glu
 20 25 30

225

Ile Ser Ala Thr Pro Phe Leu Phe Asn Pro Met His Ala Lys His Phe
 35 40 45

Ser His Leu Pro Thr His Ser Pro Ser Ala Ser Leu Ala Leu Phe Phe
 50 55 60

Thr Pro Lys Tyr Asp Arg Val Pro Ala Ala Glu Tyr Val Phe Pro Asn
 65 70 75 80

Cys Cys Gly Gln Thr Pro Val Cys Arg Ile Ala Cys Phe
 85 90

<210> 428
 <211> 59
 <212> PRT
 <213> Homo sapiens

<400> 428
 Lys Arg Ala Ser Gln Pro Pro Cys Thr Arg Asn Leu Lys Arg Ser Thr
 1 5 10 15

Asp Ser Gly Gln Arg Ala Gly Asn Ser Phe Cys Gly Asn Gln Trp Met
 20 25 30

Leu Cys Pro Thr Pro Pro His Phe Cys Trp Leu Gly Ser Pro Pro Arg
 35 40 45

Ser Thr Ser Ser Lys Arg Gly Pro Ser Ser Ser
 50 55

<210> 429
 <211> 65
 <212> PRT
 <213> Homo sapiens

<400> 429
 Pro Pro Ser Pro Pro Thr Glu Ala Ala Ser Ser Thr Ala Arg Pro Ala
 1 5 10 15

Lys Ser Arg Thr Arg Pro Thr Ser Gly Trp His Ile Gly Ser Thr Thr
 20 25 30

Pro Pro Arg Arg Ser Gln Pro Glu Val Lys Thr Leu Ala Val Asp Gln
 35 40 45

Val Asn Gly Gly Lys Val Val Arg Lys His Ser Gly Thr Asp Arg Thr
 50 55 60

Val
 65

<210> 430
 <211> 148
 <212> PRT
 <213> Homo sapiens

<400> 430
 Met Trp Asn Pro Asn Ala Gly Gln Pro Gly Pro Asn Pro Tyr Pro Pro

226

1	5	10	15
---	---	----	----

Asn Ile Gly Cys Pro Gly Gly Ser Asn Pro Ala His Pro Pro Pro Ile
 20 25 30

Asn Pro Pro Phe Pro Pro Gly Pro Cys Pro Pro Pro Pro Gly Ala Pro
 35 40 45

His Gly Asn Pro Ala Phe Pro Pro Gly Gly Pro Pro His Pro Val Pro
 50 55 60

Gln Pro Gly Tyr Pro Gly Cys Gln Pro Leu Gly Pro Tyr Pro Pro Pro
 65 70 75 80

Tyr Pro Pro Pro Ala Pro Gly Ile Pro Pro Val Asn Pro Leu Ala Pro
 85 90 95

Gly Met Val Gly Pro Ala Val Ile Val Asp Lys Lys Met Gln Lys Lys
 100 105 110

Met Lys Lys Ala His Lys Lys Met His Lys His Gln Lys His His Lys
 115 120 125

Tyr His Lys His Gly Lys His Ser Ser Ser Ser Ser Ser Ser Ser Ser
 130 135 140

Ser Asp Ser Asp
 145

<210> 431

<211> 58

<212> PRT

<213> Homo sapiens

<400> 431

Arg Val Gly Pro Asp Ala Trp Ala Asp Ala Trp Glu Gln Ala Gln Ala			
1	5	10	15

Ala Val Glu Arg Leu Glu Asp Thr Pro Lys His Val Glu Ser Gln Cys			
20	25	30	

Arg Ala Ala Arg Ala Lys Ser Ile Ser Pro Gln Tyr Trp Val Pro Trp			
35	40	45	

Arg Phe Gln Ser Cys Pro Pro Thr Thr Tyr			
50	55		

<210> 432

<211> 84

<212> PRT

<213> Homo sapiens

<400> 432

Ser Thr Leu Ser Pro Arg Pro Leu Ser Ser Ser Pro Arg Ser Ser Pro			
1	5	10	15

Trp Gln Ser Ser Phe Pro Pro Arg Trp Ala Pro Ser Ser Cys Ala Thr			
20	25	30	

227

Ala Arg Val Ser Arg Met Pro Thr Val Gly Ser Leu Pro Ser Ser Ile
 35 40 45

Pro Thr Ala Cys Pro Trp Asn Pro Ser Cys Glu Ser Leu Gly Ser Trp
 50 55 60

His Gly Trp Thr Ser Ser Asp Ser Arg Gln Glu Asp Ala Glu Glu Asn
 65 70 75 80

Glu Glu Ser Ser

<210> 433

<211> 86

<212> PRT

<213> Homo sapiens

<400> 433

Met Pro Gly Ser Gln Gly Gln Ile His Ile Pro Pro Ile Leu Gly Ala
 1 5 10 15

Leu Glu Val Pro Ile Leu Pro Thr His His Leu Leu Ile His Pro Phe
 20 25 30

Pro Gln Ala Pro Val Leu Leu Pro Gln Glu Leu Pro Met Ala Ile Gln
 35 40 45

Leu Ser Pro Gln Val Gly Pro Leu Ile Leu Cys His Ser Gln Gly Ile
 50 55 60

Gln Asp Ala Asn Arg Trp Val Pro Thr Leu Leu His Thr His Arg Leu
 65 70 75 80

Pro Leu Glu Ser Leu Leu
 85

<210> 434

<211> 65

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 434

Met Ala Ser Ile Pro Pro Leu Pro Pro Pro Leu Pro Ala Val Ile Leu
 1 5 10 15

Thr Glu Tyr Arg Pro Trp Thr Leu Pro Ser Ser Leu Thr Ser Ser Ala
 20 25 30

Leu Pro Ser Ser Phe Arg Cys His Val Val Leu Gly Glu Cys Ser Pro
 35 40 45

Cys Ala Pro His Pro Leu Pro Xaa Pro Glu Pro His Pro Ala Val Glu
 50 55 60

228

Pro

65

<210> 435

<211> 147

<212> PRT

<213> Homo sapiens

<400> 435

Pro Arg His Thr Tyr Trp Gly Ile Trp Leu Val Pro Ala Ala Met Ala
 1 5 10 15

Ser Pro His Ser His Pro Ala Gln Gly Val Leu Gln Pro Pro Gly Pro
 20 25 30

Gln Pro Arg Trp Glu Asp Arg Val Ala Leu Gly Thr Arg Gly Arg Ser
 35 40 45

Pro Gly Ala Tyr Leu Thr Glu Ser Ala Pro Gln Gln Ala Ser Thr Thr
 50 55 60

Pro Gly Pro Pro Thr Cys His Gly Lys Val Gly Ser Glu Trp Ala Trp
 65 70 75 80

Leu Gly Ala Ala Pro Gly Pro Leu Pro Thr His Pro Ser His Tyr Ala
 85 90 95

Ile Arg Val Pro Ser Asn Ile Cys Ser Cys Pro Gly Ala Ser Ser Ala
 100 105 110

Pro Ala Leu Arg Gly Val Val Arg Gln Pro Pro Gly Pro Gln Asn Pro
 115 120 125

Arg Gln Gly Gly Arg Arg Gly Thr Arg Ala Ser Pro Val Gly Ser Leu
 130 135 140

Phe Cys Val

145

<210> 436

<211> 105

<212> PRT

<213> Homo sapiens

<400> 436

Met Phe Ala Val Leu Pro Ala Val Glu Gly Arg Ala Thr Pro His Gln
 1 5 10 15

Asp Arg Thr Cys Tyr Pro Ser Arg Ser Arg Pro Trp Pro Ser Gln Pro
 20 25 30

Ser Pro Arg Gly Ser Met Pro Val Pro Arg Pro Gly Ala Ala Arg Gly
 35 40 45

Gln Leu Asp Gly His Val Gln Gly Gln Gly Trp Ala Leu Gln Trp Gly
 50 55 60

Gly Pro Pro Ala Pro Ala Val Tyr Arg Arg Met Ala Leu Pro Pro Arg
 65 70 75 80

229

Ala Ala Gly Ser Tyr Leu Asp Arg Lys Cys Pro His Pro Leu Pro Gly
 85 90 95

Ala Arg Leu Cys Pro Gly Leu Pro Leu
 100 105

<210> 437

<211> 127

<212> PRT

<213> Homo sapiens

<400> 437

Val Phe Gly Ala Val Phe Leu Thr Thr Pro Ser His Asp Leu Ala Thr
 1 5 10 15

Pro Thr Gly Ala Ser Gly Trp Cys Leu Leu Pro Trp Pro Ala Pro Thr
 20 25 30

Leu Thr Leu His Arg Gly Ser Cys Ser Pro Gln Ala His Ser Leu Val
 35 40 45

Gly Arg Thr Gly Trp Pro Trp Gly Gln Glu Gly Gly Ala Gln Gly Leu
 50 55 60

Thr Ser Leu Arg Val Leu Pro Ser Arg His Pro Leu Pro Gln Gly Pro
 65 70 75 80

Pro His Val Met Ala Arg Leu Val Val Asn Gly Pro Gly Trp Glu Gln
 85 90 95

Pro Leu Ala His Cys Pro Pro Thr His Leu Thr Met Gln Phe Glu Phe
 100 105 110

Gln Ala Thr Phe Ala Pro Ala Leu Gly Pro Ala Leu Pro Gln Pro
 115 120 125

<210> 438

<211> 186

<212> PRT

<213> Homo sapiens

<400> 438

His Glu Glu Pro Pro Ala Gly Phe Gly Leu Arg Ser Leu Trp Arg Arg
 1 5 10 15

Ser Pro Pro His Glu Val Gly Ala Arg Leu Pro Asn Gly Ala Phe Gly
 20 25 30

Phe Ser Val Arg Cys Leu Leu Cys Phe Pro Pro Trp Arg Ala Glu Pro
 35 40 45

Pro His Ile Arg Ile Gly Arg Ala Thr Pro Pro Gly Pro Gly Pro Gly
 50 55 60

Pro Ala Ser Pro Ala Leu Glu Ala Arg Cys Leu Cys Gln Gly Gln Gly
 65 70 75 80

Gln Pro Glu Gly Ser Trp Met Ala Thr Cys Arg Val Lys Ala Gly Pro

85 230 95
 90
 Cys Ser Gly Ala Gly Arg Gln Pro Gln Gln Phe Thr Asp Ala Trp Leu
 100 105 110
 Phe Leu Pro Glu Gln Pro Ala Ala Thr Trp Thr Gly Asn Val Leu Ile
 115 120 125
 Pro Ser Leu Gly Pro Gly Ser Ala Leu Ala Phe Leu Cys Glu Pro Leu
 130 135 140
 Leu Ser Leu Cys Cys Leu Gly Thr Pro Asp Arg Gly Val Arg Val Cys
 145 150 155 160
 Pro Ser Val Thr Phe Tyr Ser Pro Arg Val Glu Glu Arg Lys Arg Gly
 165 170 175
 Lys Ser Lys Gly Val Gln Thr Pro Pro Gln
 180 185

<210> 439
 <211> 100
 <212> PRT
 <213> Homo sapiens

<400> 439
 Met Ala Thr Cys Arg Val Lys Ala Gly Pro Cys Ser Gly Ala Gly Arg
 1 5 10 15
 Gln Pro Gln Gln Phe Thr Asp Ala Trp Leu Phe Leu Pro Glu Gln Pro
 20 25 30
 Ala Ala Thr Trp Thr Gly Asn Val Leu Ile Pro Ser Leu Gly Pro Gly
 35 40 45
 Ser Ala Leu Ala Phe Leu Cys Glu Pro Leu Leu Ser Leu Cys Cys Leu
 50 55 60
 Gly Thr Pro Asp Arg Gly Val Arg Val Cys Pro Ser Val Thr Phe Tyr
 65 70 75 80
 Ser Pro Arg Val Glu Glu Arg Lys Arg Gly Lys Ser Lys Gly Val Gln
 85 90 95
 Thr Pro Pro Gln
 100

<210> 440
 <211> 244
 <212> PRT
 <213> Homo sapiens

<400> 440
 Met Lys Trp Phe Ser Thr Gln Pro Leu Trp Leu Asn Thr Lys Gln Arg
 1 5 10 15
 Ser His Arg Arg Gly Pro Gly Pro Pro Pro Ala Pro Leu Ser Gly Val
 20 25 30

231

Leu Gly Ser Arg Gly Leu Pro His His Pro Ser Gln Gly Trp Gly Arg
 35 40 45
 Ala Gly Pro Arg Ala Gly Ala Asn Val Ala Trp Asn Ser Asn Cys Ile
 50 55 60
 Val Arg Trp Val Gly Gly Gln Trp Ala Arg Gly Cys Ser Gln Pro Gly
 65 70 75 80
 Pro Phe Thr Thr Asn Leu Ala Met Thr Cys Gly Gly Pro Trp Gly Ser
 85 90 95
 Gly Cys Leu Leu Gly Ser Thr Leu Ser Glu Val Ser Pro Trp Ala Pro
 100 105 110
 Pro Ser Cys Pro Gln Gly His Pro Val Leu Pro Thr Arg Leu Trp Ala
 115 120 125
 Trp Gly Leu Gln Asp Pro Leu Cys Arg Val Arg Val Gly Ala Gly His
 130 135 140
 Gly Ser Arg His Gln Pro Asp Ala Pro Val Gly Val Ala Arg Ser Trp
 145 150 155 160
 Asp Gly Val Val Arg Asn Thr Ala Pro Lys Thr Gln Asn Lys Asn Thr
 165 170 175
 Thr Asn Gly Arg Arg Ser Pro Pro Pro Thr Glu Val Gly Phe Glu Pro
 180 185 190
 Leu Leu Ile Phe Pro Val Ser Phe Leu Gln Pro Leu Val Ser Arg Lys
 195 200 205
 Ser Gln Thr Gly Thr His Ala His His Gly Gln Glu Ser Arg Asp Ser
 210 215 220
 Thr Lys Lys Gly Gly Val His Arg Gly Arg Pro Gly Gln Ser Leu Ala
 225 230 235 240
 Pro Gly Arg Gly

<210> 441

<211> 165

<212> PRT

<213> Homo sapiens

<400> 441

Lys Val Thr Asp Gly His Thr Arg Thr Pro Arg Ser Gly Val Pro Arg
 1 5 10 15
 Gln His Lys Glu Arg Arg Gly Ser Gln Arg Lys Ala Arg Ala Glu Pro
 20 25 30
 Gly Pro Arg Glu Gly Met Arg Thr Phe Pro Val Gln Val Ala Ala Gly
 35 40 45
 Cys Ser Gly Arg Lys Ser His Ala Ser Val Asn Cys Trp Gly Trp Arg
 50 55 60

232

Pro Ala Pro Leu Gln Gly Pro Ala Leu Thr Leu His Val Ala Ile Gln
 65 70 75 80

Leu Pro Ser Gly Cys Pro Trp Pro Trp His Arg His Arg Ala Ser Arg
 85 90 95

Ala Gly Leu Ala Gly Pro Gly Pro Gly Pro Gly Gly Val Ala Arg Pro
 100 105 110

Ile Leu Met Trp Gly Gly Ser Ala Leu His Gly Gly Lys His Ser Lys
 115 120 125

His Arg Thr Leu Lys Pro Lys Ala Pro Leu Gly Ser Leu Ala Pro Thr
 130 135 140

Ser Trp Gly Gly Asp Arg Arg His Arg Asp Leu Ser Pro Lys Pro Ala
 145 150 155 160

Gly Gly Ser Ser Cys
 165

<210> 442

<211> 128

<212> PRT

<213> Homo sapiens

<400> 442

Met Arg Thr Phe Pro Val Gln Val Ala Ala Gly Cys Ser Gly Arg Lys
 1 5 10 15

Ser His Ala Ser Val Asn Cys Trp Gly Trp Arg Pro Ala Pro Leu Gln
 20 25 30

Gly Pro Ala Leu Thr Leu His Val Ala Ile Gln Leu Pro Ser Gly Cys
 35 40 45

Pro Trp Pro Trp His Arg His Arg Ala Ser Arg Ala Gly Leu Ala Gly
 50 55 60

Pro Gly Pro Gly Pro Gly Gly Val Ala Arg Pro Ile Leu Met Trp Gly
 65 70 75 80

Gly Ser Ala Leu His Gly Gly Lys His Ser Lys His Arg Thr Leu Lys
 85 90 95

Pro Lys Ala Pro Leu Gly Ser Leu Ala Pro Thr Ser Trp Gly Gly Asp
 100 105 110

Arg Arg His Arg Asp Leu Ser Pro Lys Pro Ala Gly Gly Ser Ser Cys
 115 120 125

<210> 443

<211> 13

<212> PRT

<213> Homo sapiens

233

<400> 443

Gly Leu Met Glu Cys Leu Ile His Arg His Gly Ser His
1 5 10

<210> 444

<211> 17

<212> PRT

<213> Homo sapiens

<400> 444

Ser Thr Lys Gly Met Gln Phe Ile Leu Thr Gly Ile Thr Leu Ser Gly
1 5 10 15

Tyr

<210> 445

<211> 209

<212> PRT

<213> Homo sapiens

<400> 445

Pro Arg Val Arg Ala Leu Leu Phe Ala Arg Ser Leu Arg Leu Cys Arg
1 5 10 15

Trp Gly Ala Lys Arg Leu Gly Val Ala Ser Thr Glu Ala Gln Arg Gly
20 25 30

Val Ser Phe Lys Leu Glu Glu Lys Thr Ala His Ser Ser Leu Ala Leu
35 40 45

Phe Arg Asp Asp Thr Gly Val Lys Tyr Gly Leu Val Gly Leu Glu Pro
50 55 60

Thr Lys Val Ala Leu Asn Val Glu Arg Phe Arg Glu Trp Ala Val Val
65 70 75 80

Leu Ala Asp Thr Ala Val Thr Ser Gly Arg His Tyr Trp Glu Val Thr
85 90 95

Val Lys Arg Ser Gln Gln Phe Arg Ile Gly Val Ala Asp Val Asp Met
100 105 110

Ser Arg Asp Ser Cys Ile Gly Val Asp Asp Arg Ser Trp Val Phe Thr
115 120 125

Met Pro Ser Ala Ser Gly Thr Pro Cys Trp Pro Thr Arg Lys Pro Gln
130 135 140

Leu Arg Val Leu Gly Ser Gln Glu Val Gly Leu Leu Leu Glu Tyr Glu
145 150 155 160

Ala Gln Lys Leu Ser Leu Val Asp Val Ser Gln Val Ser Val Val His
165 170 175

Thr Leu Gln Thr Asp Phe Arg Gly Pro Val Val Pro Ala Phe Ala Leu
180 185 190

234

Trp Asp Gly Glu Leu Leu Thr His Ser Gly Leu Glu Val Pro Glu Gly
 195 200 205

Leu

<210> 446
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 446
 Met Ser Arg Asp Ser Cys Ile Gly Val Asp Asp Arg Ser Trp Val Phe
 1 5 10 15

Thr Met Pro Ser Ala Ser Gly Thr Pro Cys Trp Pro Thr Arg Lys Pro
 20 25 30

Gln Leu Arg Val Leu Gly Ser Gln Glu Val Gly Leu Leu Leu Glu Tyr
 35 40 45

Glu Ala Gln Lys Leu Ser Leu Val Asp Val Ser Gln Val Ser Val Val
 50 55 60

His Thr Leu Gln Thr Asp Phe Arg Gly Pro Val Val Pro Ala Phe Ala
 65 70 75 80

Leu Trp Asp Gly Glu Leu Leu Thr His Ser Gly Leu Glu Val Pro Glu
 85 90 95

Gly Leu

<210> 447
 <211> 1913
 <212> DNA
 <213> Homo sapiens

<400> 447
 GCACGAGCGG CACGAGCGGA TCCTCACACG ACTGTGATCC GATTCTTTCC AGCGGCTTCT 60
 GCAACCAAGC GGGTCTTACC CCCGGTCCTC CGCGTCTCCA GTCCTCGCAC CTGGAACCCC 120
 AACGTCCCCG AGAGTCCCCG AATCCCCGCT CCCAGGCTAC CTAAGAGGAT GAGCGGTGCT 180
 CCGACGGCCG GGGCAGCCCT GATGCTCTGC GCCGCCACCG CCGTGCTACT GAGCGCTCAG 240
 GGCGGACCCG TGCAGTCCAA GTCGCCGCGC TTTGCGTCCT GGGACGAGAT GAATGTCCTG 300
 GCGCACGGAC TCCTGCAGCT CGGCCAGGGG CTGCGCGAAC ACGCGGAGCG CACCCGCACT 360
 CAGCTGAGCG CGCTGGAGCG GCGCCTGAGC GCGTGCGGGT CCGCCTGTCA GGAACCCGAG 420
 GGGTCCACCG ACCTCCCGTT AGCCCCTGAG AGCCGGGTGG ACCCTGAGGT CCTTCACAGC 480
 CTGCAGACAC AACTCAAGGC TCAGAACAGC AGGATCCAGC AACTCTTCCA CAAGGTGGCC 540
 CAGCAGCAGC GGCACCTGGA GAAGCAGCAC CTGCGAATTC AGCATCTGCA AAGCCAGTTT 600

235

GGCCTCCTGG ACCACAAGCA CCTAGACCAT GAGGTGGCCA AGCCTGCCCG AAGAAAGAGG 660
CTGCCCAGAGA TGGCCCAGCC AGTTGACCCG GCTCACAATG TCAGCCGCCT GCACCGGCTG 720
CCCAGGGATT GCCAGGAGCT GTTCCAGGTT GGGGAGAGGC AGAGTGGACT ATTTGAAATC 780
CAGCCTCAGG GGTCTCCGCC ATTTTGGTG AACTGCAAGA TGACCTCAGA TGGAGGCTGG 840
ACAGTAATTC AGAGGCGCCA CGATGGCTCA GTGGACTTCA ACCGGCCCTG GGAAGCCTAC 900
AAGGCGGGGT TTGGGGATCC CCACGGCGAG TTCTGGCTGG GTCTGGAGAA GGTGCATAGC 960
ATCACGGGGG ACCGCAACAG CCGCCTGGCC GTGCAGCTGC GGGACTGGGA TGGCAACGCC 1020
GAGTTGCTGC AGTTCTCCGT GCACCTGGGT GCGGAGGACA CGGCCTATAG CCTGCAGCTC 1080
ACTGCACCCG TGGCCGGCCA GCTGGGCGCC ACCACCGTCC CACCCAGCGG CCTCTCCGTA 1140
CCCTTCTCCA CTGGGACCA GGATCACGAC CTCCGCAGGG ACAAGAACTG CGCCAAGAGC 1200
CTCTCTGGAG GCTGGTGGTT TGGCACCTGC AGCCATTCCA ACCTCAACGG CCAGTACTTC 1260
CGCTCCATCC CACAGCAGCG GCAGAAGCTT AAGAAGGGAA TCTTCTGGAA GACCTGGCGG 1320
GGCCGCTACT ACCCGCTGCA GGCCACCACC ATGTTGATCC AGCCCATGGC AGCAGAGGCA 1380
GCCTCCTAGC GTCCTGGCTG GGCCTGGTCC CAGGCCACG AAAGACGGTG ACTCTTGGCT 1440
CTGCCCAGAG ATGTGGCCGT TCCCTGCCTG GGCAGGGGCT CCAAGGAGGG GCCATCTGGA 1500
AACTTGTTGA CAGAGAAGAA GACCACGACT GGAGAAGCCC CCTTTCTGAG TGCAGGGGGG 1560
CTGCATGCGT TGCCTCCTGA GATCGAGGCT GCAGGATATG CTCAGACTCT AGAGGCGTGG 1620
ACCAAGGGGC ATGGAGCTTC ACTCCTTGCT GGCCAGGGAG TTGGGGACTC AGAGGGACCA 1680
CTTGGGGCCA GCCAGACTGG CCTCAATGGC GGACTCAGTC ACATTGACTG ACGGGGACCA 1740
GGGCTTGTGT GGGTCGAGAG CGCCCTCATG GTGCTGGTGC TGTGTGTGT AGGTCCCCTG 1800
GGGACACAAG CAGGCGCCAA TGGTATCTGG GCGGAGCTCA CAGAGTTCTT GGAATAAAAG 1860
CAACCTCAGA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA 1913

<210> 448

<211> 1221

<212> DNA

<213> Homo sapiens

<400> 448

ATGAGCGGTG CTCCGACGGC CGGGGAGGCC CTGATGCTCT GCGCCGCCAC CGCCGTGCTA 60
CTGAGCGCTC AGGGCGGACC CGTGCACTCC AAGTCGCCGC GCTTTGCGTC CTGGGACGAG 120
ATGAATGTCC TGGCGCACGG ACTCCTGCAG CTCGGCCAGG GGCTGCGCGA ACACGCGGAG 180
CGCACCCGCA GTCAGCTGAG CGCGCTGGAG CGGCGCCTGA GCGCGTGC GG GTCCGCCTGT 240

236

CAGGGAACCG AGGGGTCCAC CGACCTCCCG TTAGCCCCTG AGAGCCGGGT GGACCCTGAG	300
GTCCTTCACA GCCTGCAGAC ACAACTCAAG GCTCAGAACA GCAGGATCCA GCAACTCTTC	360
CACAAGGTGG CCCAGCAGCA GCGGCACCTG GAGAAGCAGC ACCTGCGAAT TCAGCATCTG	420
CAAAGCCAGT TTGGCCTCCT GGACCACAAG CACCTAGACC ATGAGGTGGC CAAGCCTGCC	480
CGAAGAAAGA GGCTGCCCCG GATGGCCCAG CCAGTTGACC CGGCTCACAA TGTCAGCCGC	540
CTGCACCGGC TGCCCAGGGA TTGCCAGGAG CTGTTCCAGG TTGGGGAGAG GCAGAGTGA	600
CTATTTGAAA TCCAGCCTCA GGGGTCTCCG CCATTTTGG TGAAGTCAA GATGACCTCA	660
GATGGAGGCT GGACAGTAAT TCAGAGGCGC CACGATGGCT CAGTGGACTT CAACCGCCCC	720
TGGGAAGCCT ACAAGGCGGG GTTTGGGGAT CCCACGGCG AGTTCTGGCT GGGTCTGGAG	780
AAGGTGCATA GCATCACGGG GGACCGCAAC AGCCGCCTGG CCGTGCAGCT GCGGGACTGG	840
GATGGCAACG CCGAGTTGCT GCAGTTCTCC GTGCACCTGG GTGGCGAGGA CACGGCCTAT	900
AGCCTGCAGC TCACTGCACC CGTGGCCGGC CAGCTGGGCG CCACCACCGT CCCACCCAGC	960
GGCCTCTCCG TACCCTTCTC CACTTGGGAC CAGGATCACG ACCTCCGCAG GGACAAGAAC	1020
TGCGCCAAGA GCCTCTCTGG AGGCTGGTGG TTTGGCACCT GCAGCCATTC CAACCTCAAC	1080
GGCCAGTACT TCCGCTCCAT CCCACAGCAG CGGCAGAAGC TTAAGAAGGG AATCTTCTGG	1140
AAGACCTGGC GGGGCCGCTA CTACCCGCTG CAGGCCACCA CCATGTTGAT CCAGCCCATG	1200
GCAGCAGAGG CAGCCTCCTA G	1221

<210> 449

<211> 175

<212> PRT

<213> Homo sapiens

<400> 449

Met Ala Gln Trp Thr Ser Thr Gly Pro Gly Lys Pro Thr Arg Arg Gly
 1 5 10 15

Leu Gly Ile Pro Thr Ala Ser Ser Gly Trp Val Trp Arg Arg Cys Ile
 20 25 30

Ala Ser Trp Gly Thr Ala Thr Ala Ala Trp Pro Cys Ser Cys Gly Thr
 35 40 45

Gly Met Ala Thr Pro Ser Cys Cys Ser Ser Pro Cys Thr Trp Val Ala
 50 55 60

Arg Thr Arg Pro Ile Ala Cys Ser Ser Leu His Pro Trp Pro Ala Ser
 65 70 75 80

Trp Ala Pro Pro Pro Ser His Pro Ala Ala Ser Pro Tyr Pro Ser Pro
 85 90 95

237

Leu Gly Thr Arg Ile Thr Thr Ser Ala Gly Thr Arg Thr Ala Pro Arg
 100 105 110

Ala Ser Leu Glu Ala Gly Gly Leu Ala Pro Ala Ala Ile Pro Thr Phe
 115 120 125

Asn Gly Pro Val Leu Pro Ala Pro Ser His Ser Ser Gly Arg Ser Leu
 130 135 140

Arg Arg Glu Ser Ser Gly Arg Pro Ala Gly Arg Tyr Tyr Pro Leu Gln
 145 150 155 160

Ala Thr Thr Met Leu Ile Gln Pro Met Ala Ala Glu Ala Ala Ser
 165 170 175

<210> 450

<211> 32

<212> PRT

<213> Homo sapiens

<400> 450

Gly His Asp Leu Pro Gln Asp Ala Trp Leu Arg Trp Val Leu Ala Gly
 1 5 10 15

Ala Leu Cys Ala Gly Gly Trp Ala Val Asn Tyr Leu Pro Phe Phe Leu
 20 25 30

<210> 451

<211> 18

<212> PRT

<213> Homo sapiens

<400> 451

Phe Leu Tyr His Tyr Leu Pro Ala Leu Thr Phe Gln Ile Leu Leu Leu
 1 5 10 15

Pro Val

<210> 452

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (44)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (49)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 452

Met Ser Pro Leu Pro Trp Pro Gly Pro Leu Pro Gly Gly Arg Gln Gly

1 5 10 15

His Arg Leu Glu Pro Cys Cys Ser Ser Gly Cys Ala Gly Gly Pro Thr
20 25 30

Trp Pro His Cys Ser Ser Gln Ser Trp Pro Met Xaa Ser Ala Arg His
35 40 45

Xaa Gly Leu Gly His Cys Cys Pro Ser Ser Pro
50 55

<210> 453
<211> 32
<212> PRT
<213> Homo sapiens

<400> 453
Asp Ile Cys Arg Leu Glu Arg Ala Val Cys Arg Asp Glu Pro Ser Ala
1 5 10 15

Leu Ala Arg Ala Leu Thr Trp Arg Gln Ala Arg Ala Gln Ala Gly Ala
20 25 30

<210> 454
<211> 114
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (1)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (6)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 454
Xaa Ala Pro Ala Thr Xaa Ala Trp Asp Thr Val Val Pro Pro Leu Pro
1 5 10 15

Arg Lys Cys Gln Cys Ser Gly Ser Ala Arg Ser His Gly Ala Gly Arg
20 25 30

Ser Ala Leu His Ser Pro Leu Glu Gly Ser Arg Pro Lys Val Pro Ala
35 40 45

Gly Ala Val Gly Lys Ser Leu Pro Gly Gln Ser Arg Pro Gln His Cys
50 55 60

Leu Pro Pro Lys Gln Pro Lys Gln Cys Arg Pro Gly Leu Glu Leu Lys
65 70 75 80

Glu Gly Pro Leu Leu Thr Pro Thr Arg Ala Ser Val Gln Leu Ser His
85 90 95

Pro Ala Cys Leu Tyr Trp Ala Pro Leu Leu Trp Ile Arg Asp Pro Ala
 100 105 110

Ser Val

<210> 455

<211> 55

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (1)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (6)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 455

Xaa Ala Pro Ala Thr Xaa Ala Trp Asp Thr Val Val Pro Pro Leu Pro
 1 5 10 15

Arg Lys Cys Gln Cys Ser Gly Ser Ala Arg Ser His Gly Ala Gly Arg
 20 25 30

Ser Ala Leu His Ser Pro Leu Glu Gly Ser Arg Pro Lys Val Pro Ala
 35 40 45

Gly Ala Val Gly Lys Ser Leu
 50 55

<210> 456

<211> 59

<212> PRT

<213> Homo sapiens

<400> 456

Pro Gly Gln Ser Arg Pro Gln His Cys Leu Pro Pro Lys Gln Pro Lys
 1 5 10 15

Gln Cys Arg Pro Gly Leu Glu Leu Lys Glu Gly Pro Leu Leu Thr Pro
 20 25 30

Thr Arg Ala Ser Val Gln Leu Ser His Pro Ala Cys Leu Tyr Trp Ala
 35 40 45

Pro Leu Leu Trp Ile Arg Asp Pro Ala Ser Val
 50 55

<210> 457

<211> 133

<212> PRT

<213> Homo sapiens

<220>

240

<221> SITE

<222> (55)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (61)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 457

Asp	Ile	Cys	Arg	Leu	Glu	Arg	Ala	Val	Cys	Arg	Asp	Glu	Pro	Ser	Ala
1				5					10					15	

Leu	Ala	Arg	Ala	Leu	Thr	Trp	Arg	Gln	Ala	Arg	Ala	Gln	Ala	Gly	Ala
			20					25					30		

Met	Leu	Leu	Phe	Gly	Leu	Cys	Trp	Gly	Pro	Tyr	Val	Ala	Thr	Leu	Leu
			35				40					45			

Leu	Ser	Val	Leu	Ala	Tyr	Xaa	Gln	Arg	Pro	Pro	Leu	Xaa	Pro	Gly	Thr
	50					55					60				

Leu	Leu	Ser	Leu	Leu	Ser	Leu	Gly	Ser	Ala	Ser	Ala	Ala	Ala	Val	Pro
65					70					75				80	

Val	Ala	Met	Gly	Leu	Gly	Asp	Gln	Arg	Tyr	Thr	Ala	Pro	Trp	Arg	Ala
				85					90					95	

Ala	Ala	Gln	Arg	Cys	Leu	Gln	Gly	Leu	Trp	Gly	Arg	Ala	Ser	Arg	Asp
			100					105					110		

Ser	Pro	Gly	Pro	Ser	Ile	Ala	Tyr	His	Pro	Ser	Ser	Gln	Ser	Ser	Val
		115					120					125			

Asp	Leu	Asp	Leu	Asn
	130			

<210> 458

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (34)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (43)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 458

Met	Glu	Arg	Val	Gly	Met	Glu	Ser	Gly	Glu	Met	Val	Cys	Gly	Leu	Gly
1				5					10					15	

Ser	Ala	Cys	Asn	Asn	Pro	Ser	Asp	Leu	Gly	Gln	Val	Pro	Val	Pro	Leu
			20					25					30		

241

Trp Xaa Ser Val Ser Pro Pro Val Phe Gly Xaa Gly Trp Asn Gly His
35 40 45

<210> 459

<211> 107

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (84)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 459

Met Arg Ser Phe Gln Asp Val Ser Ala Leu Glu Glu Trp Arg Gly Gly
1 5 10 15

Lys Asp Leu Glu Pro Thr His Ser Leu Leu Leu Leu Leu Pro Leu Arg
20 25 30

Asp Leu Leu Val Val Leu Gly Glu Ile Arg Lys Arg Gln Met Glu Gly
35 40 45

Cys Val Trp Lys Gly Trp Gly Trp Asn Pro Glu Lys Trp Phe Ala Val
50 55 60

Leu Ala Leu Pro Val Thr Thr Arg Val Thr Leu Gly Lys Ser Leu Ser
65 70 75 80

Leu Ser Gly Xaa Gln Phe Leu His Leu Tyr Leu Glu Arg Val Gly Met
85 90 95

Gly Thr Glu Val Leu Ser Ser Ser Asp Leu Leu
100 105

<210> 460

<211> 118

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (62)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (70)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 460

Met His Pro Ala Gly Pro Thr Phe Met Gly Ser Lys Pro Ile Arg Glu
1 5 10 15

Gln Gln Phe Gly Pro Asp Ala Cys Leu Leu Leu Leu Cys Val Ala Met
20 25 30

242

Ala Gly Thr Glu Ala Ser Arg Ala Ala Gln Gln Cys Thr Ser Gln Lys
 35 40 45

Val Arg Ala Gly Gln Asp Phe Ser Ala His Ser Asn Pro Xaa Gln Ile
 50 55 60

Gln Val Glu Lys Leu Xaa Pro Arg Glu Gly Gln Gly Leu Ala Gln Gly
 65 70 75 80

His Ser Gly Cys Tyr Arg Gln Ser Gln Asp Arg Lys Pro Phe Leu Arg
 85 90 95

Ile Pro Ser Pro Pro Phe Pro Tyr Thr Thr Leu His Leu Pro Phe Pro
 100 105 110

Asp Phe Ala Lys Asn His
 115

<210> 461
 <211> 61
 <212> PRT
 <213> Homo sapiens

<400> 461
 Met His Pro Ala Gly Pro Thr Phe Met Gly Ser Lys Pro Ile Arg Glu
 1 5 10 15

Gln Gln Phe Gly Pro Asp Ala Cys Leu Leu Leu Cys Val Ala Met
 20 25 30

Ala Gly Thr Glu Ala Ser Arg Ala Ala Gln Gln Cys Thr Ser Gln Lys
 35 40 45

Val Arg Ala Gly Gln Asp Phe Ser Ala His Ser Asn Pro
 50 55 60

<210> 462
 <211> 48
 <212> PRT
 <213> Homo sapiens

<400> 462
 Pro Arg Glu Gly Gln Gly Leu Ala Gln Gly His Ser Gly Cys Tyr Arg
 1 5 10 15

Gln Ser Gln Asp Arg Lys Pro Phe Leu Arg Ile Pro Ser Pro Pro Phe
 20 25 30

Pro Tyr Thr Thr Leu His Leu Pro Phe Pro Asp Phe Ala Lys Asn His
 35 40 45

<210> 463
 <211> 22
 <212> PRT
 <213> Homo sapiens

243

<400> 463

Asp Pro Arg Val Arg Lys Pro Pro Thr Ala Thr Leu Thr Thr Ala Arg
 1 5 10 15

Thr Arg Pro Thr Thr Asp
 20

<210> 464

<211> 82

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (70)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (81)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (82)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 464

Ala Ala Leu Glu Ala Ser Val Pro Ala Ile Ala Thr Gln Arg Ser Ser
 1 5 10 15

Arg Gln Ala Ser Gly Pro Asn Cys Cys Ser Leu Met Gly Leu Asp Pro
 20 25 30

Met Lys Val Gly Pro Ala Gly Cys Ile Ser Trp Asp Ser Val Glu Ala
 35 40 45

Asp Gln Val Ala Gly Ala Ser Gly Gly Arg Ile Glu Val Lys Gly Cys
 50 55 60

Gly Met Glu Asn Leu Xaa Arg Leu His Leu Gly Ser Gly Lys Gly Gln
 65 70 75 80

Xaa Xaa

<210> 465

<211> 99

<212> PRT

<213> Homo sapiens

<400> 465

Met Leu His Arg Gln Trp Leu Thr Val Arg Arg Ala Gly Gly Pro Pro
 1 5 10 15

Arg Thr Asp Gln Gln Arg Arg Thr Val Arg Cys Leu Arg Asp Thr Val
 20 25 30

244

Leu Leu Leu His Gly Leu Ser Gln Lys Asp Lys Leu Phe Met Met His
 35 40 45

Cys Val Glu Val Leu His Gln Phe Asp Gln Val Met Pro Gly Val Ser
 50 55 60

Met Leu Ile Arg Gly Leu Pro Asp Val Thr Asp Cys Glu Glu Ala Ala
 65 70 75 80

Leu Asp Asp Leu Cys Ala Ala Glu Thr Asp Val Glu Asp Pro Glu Val
 85 90 95

Glu Cys Gly

<210> 466

<211> 62

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (2)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (58)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 466

Gly Xaa Ala Asn Pro Glu Asp Ser Val Cys Ile Leu Glu Gly Phe Ser
 1 5 10 15

Val Thr Ala Leu Ser Ile Leu Gln His Leu Val Cys His Ser Gly Ala
 20 25 30

Val Arg Leu Pro Ile Thr Val Arg Ser Gly Gly Arg Phe Cys Cys Trp
 35 40 45

Gly Arg Lys Gln Glu Pro Gly Ser Gln Xaa Ser Asp Gly Asp
 50 55 60

<210> 467

<211> 65

<212> PRT

<213> Homo sapiens

<400> 467

Ala Val Gln Gln Gln His Arg Val Pro Gln Thr Ala His Cys Pro Pro
 1 5 10 15

Leu Leu Val Gly Pro Trp Gly Ser Pro Cys Pro Pro His Cys Gln Pro
 20 25 30

Leu Ser Val Gln His His Arg Glu Arg Ser Asp His Leu His Ile Thr
 35 40 45

Leu Ala Val Gly Ala Ser Asp Trp Gly Gln Gly Ala Leu Ala His Gln

245

50

55

60

Ala

65

<210> 468

<211> 220

<212> PRT

<213> Homo sapiens

<400> 468

Pro Lys Thr Leu Pro Val Ile Ser Cys Pro Gly Ser Ser Val Cys Ser
 1 5 10 15

Lys Cys Cys Gln Ser Ala Ser Ala Gln Arg His Pro Cys Leu Ala Cys
 20 25 30

Cys Trp Leu Leu Ser Ser Ser Pro Cys Trp Arg Thr Thr Thr Ser Trp
 35 40 45

His Leu Ser Ser Val Pro Thr Gln Lys Ala Ala Ser Cys Cys Cys Cys
 50 55 60

Thr Cys Thr Ser His His Gly Leu Thr Glu Trp Pro Trp Arg His Asn
 65 70 75 80

Gly Ser Ser Trp Asn Lys Arg Trp Cys Gly Ser Trp Leu Ser Leu Val
 85 90 95

Cys Lys Ser Pro Leu Pro Pro Val Thr Gly Ser Asn Cys Gln Cys Asn
 100 105 110

Val Glu Val Val Arg Ala Leu Thr Val Met Leu His Arg Gln Trp Leu
 115 120 125

Thr Val Arg Arg Ala Gly Gly Pro Pro Arg Thr Asp Gln Gln Arg Arg
 130 135 140

Thr Val Arg Cys Leu Arg Asp Thr Val Leu Leu Leu His Gly Leu Ser
 145 150 155 160

Gln Lys Asp Lys Leu Phe Met Met His Cys Val Glu Val Leu His Gln
 165 170 175

Phe Asp Gln Val Met Pro Gly Val Ser Met Leu Ile Arg Gly Leu Pro
 180 185 190

Asp Val Thr Asp Cys Glu Glu Ala Ala Leu Asp Asp Leu Cys Ala Ala
 195 200 205

Glu Thr Asp Val Glu Asp Pro Glu Val Glu Cys Gly
 210 215 220

<210> 469

<211> 223

<212> PRT

<213> Homo sapiens

<220>

246

<221> SITE

<222> (2)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (58)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 469

Gly	Xaa	Ala	Asn	Pro	Glu	Asp	Ser	Val	Cys	Ile	Leu	Glu	Gly	Phe	Ser
1				5					10					15	

Val	Thr	Ala	Leu	Ser	Ile	Leu	Gln	His	Leu	Val	Cys	His	Ser	Gly	Ala
			20					25					30		

Val	Arg	Leu	Pro	Ile	Thr	Val	Arg	Ser	Gly	Gly	Arg	Phe	Cys	Cys	Trp
	35						40					45			

Gly	Arg	Lys	Gln	Glu	Pro	Gly	Ser	Gln	Xaa	Ser	Asp	Gly	Asp	Met	Thr
	50					55					60				

Ser	Ala	Leu	Arg	Gly	Val	Ala	Asp	Asp	Gln	Gly	Gln	His	Pro	Leu	Leu
65					70					75					80

Lys	Met	Leu	Leu	His	Leu	Leu	Ala	Phe	Ser	Ser	Ala	Ala	Thr	Gly	His
				85					90					95	

Leu	Gln	Ala	Ser	Val	Leu	Thr	Gln	Cys	Leu	Lys	Val	Leu	Val	Lys	Leu
			100					105					110		

Ala	Glu	Asn	Thr	Ser	Cys	Asp	Phe	Leu	Pro	Arg	Phe	Gln	Cys	Val	Phe
		115					120					125			

Gln	Val	Leu	Pro	Lys	Cys	Leu	Ser	Pro	Glu	Thr	Pro	Leu	Pro	Ser	Val
	130					135					140				

Leu	Leu	Ala	Val	Glu	Leu	Leu	Ser	Leu	Leu	Ala	Asp	His	Asp	Gln	Leu
145					150					155				160	

Ala	Pro	Gln	Leu	Cys	Ser	His	Ser	Glu	Gly	Cys	Leu	Leu	Leu	Leu	Leu
			165						170					175	

Tyr	Met	Tyr	Ile	Thr	Ser	Arg	Pro	Asp	Arg	Val	Ala	Leu	Glu	Thr	Gln
			180					185					190		

Trp	Leu	Gln	Leu	Glu	Gln	Glu	Val	Val	Trp	Leu	Leu	Ala	Lys	Leu	Gly
	195					200						205			

Val	Gln	Glu	Pro	Leu	Ala	Pro	Ser	His	Trp	Leu	Gln	Leu	Pro	Val
	210					215					220			

<210> 470

<211> 102

<212> PRT

<213> Homo sapiens

<400> 470

Met Ser Gly Gln Leu Asp Ala Arg Pro Ala Ala Ala Leu His Pro Gln

1 5 247 15
 10
 Gly Leu Ala His Pro Leu Trp Thr Cys Leu Leu Pro Arg Lys Gly Pro
 20 25 30
 Ser Glu Val Pro Gln Arg Pro Pro Gln Leu Trp Val Val Ser Ile Ser
 35 40 45
 Val Leu Gln Gly Gln His Arg Gly Arg Ala Gly Pro Arg Asp Glu Gln
 50 55 60
 Ser Val Asp Val Thr Asn Thr Thr Phe Leu Leu Met Ala Ala Ser Ile
 65 70 75 80
 Tyr Leu His Asp Gln Asn Pro Asp Ala Ala Leu Arg Ala Leu His Gln
 85 90 95
 Gly Asp Ser Leu Glu Trp
 100

<210> 471
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 471
 Ser Val Asp Val Thr Asn Thr Thr Phe Leu Leu Met Ala Ala Ser Ile
 1 5 10 15

Tyr Leu His Asp
 20

<210> 472
 <211> 17
 <212> PRT
 <213> Homo sapiens

<400> 472
 Gln Asn Pro Asp Ala Ala Leu Arg Ala Leu His Gln Gly Asp Ser Leu
 1 5 10 15

Glu

<210> 473
 <211> 14
 <212> PRT
 <213> Homo sapiens

<400> 473
 Arg Asp Ser Ile Val Ala Glu Leu Asp Arg Glu Met Ser Arg
 1 5 10

<210> 474
 <211> 39
 <212> PRT
 <213> Homo sapiens

<400> 474

248

Met Leu Gly Leu Leu Leu Leu Cys Thr Pro Arg Ala Trp Leu Thr Leu
1 5 10 15

Ser Gly Pro Val Cys Phe Gln Gly Arg Asp Pro Leu Arg Ser His Arg
20 25 30

Gly His Pro Ser Cys Gly Ser
35

<210> 475

<211> 11

<212> PRT

<213> Homo sapiens

<400> 475

His Gly Phe Pro Glu Phe Trp Tyr Ser Trp Arg
1 5 10

<210> 476

<211> 10

<212> PRT

<213> Homo sapiens

<400> 476

Ala Ser His Trp Leu Gln Gln Asp Gln Pro
1 5 10

<210> 477

<211> 9

<212> PRT

<213> Homo sapiens

<400> 477

Pro Ile Asn His Tyr Arg Asn Ile Phe
1 5

<210> 478

<211> 9

<212> PRT

<213> Homo sapiens

<400> 478

Tyr Pro Glu Met Val Met Lys Leu Ile
1 5

<210> 479

<211> 14

<212> PRT

<213> Homo sapiens

<400> 479

Pro Glu Phe Trp Tyr Ser Trp Arg Tyr Gln Leu Arg Glu Phe
1 5 10

<210> 480

<211> 9

<212> PRT

<213> Homo sapiens

249

<400> 480

His Asp Trp Gly Gly Met Ile Ala Trp
1 5

<210> 481

<211> 14

<212> PRT

<213> Homo sapiens

<400> 481

Gly Ser Leu Pro Pro Lys Pro Ile Tyr Leu Val Val Pro Arg
1 5 10

<210> 482

<211> 16

<212> PRT

<213> Homo sapiens

<400> 482

Leu Val Phe Ala Glu His Arg Tyr Tyr Gly Lys Ser Leu Pro Phe Gly
1 5 10 15

<210> 483

<211> 10

<212> PRT

<213> Homo sapiens

<400> 483

Glu Gln Ala Leu Ala Asp Phe Ala Glu Leu
1 5 10

<210> 484

<211> 18

<212> PRT

<213> Homo sapiens

<400> 484

Gly Gly Ser Tyr Gly Gly Met Leu Ser Ala Tyr Leu Arg Met Lys Tyr
1 5 10 15

Pro His

<210> 485

<211> 16

<212> PRT

<213> Homo sapiens

<400> 485

Asn Ile Ile Phe Ser Asn Gly Asn Leu Asp Pro Trp Ala Gly Gly Gly
1 5 10 15

<210> 486

250

<211> 22

<212> PRT

<213> Homo sapiens

<400> 486

Ala Met Met Asp Tyr Pro Tyr Pro Thr Asp Phe Leu Gly Pro Leu Pro
1 5 10 15

Ala Asn Pro Val Lys Val
20

<210> 487

<211> 8

<212> PRT

<213> Homo sapiens

<400> 487

Phe Tyr Thr Gly Asn Glu Gly Asp
1 5

<210> 488

<211> 490

<212> PRT

<213> Homo sapiens

<400> 488

Met Gly Ser Ala Pro Trp Ala Pro Val Leu Leu Ala Leu Gly Leu
1 5 10 15

Arg Gly Leu Gln Ala Gly Ala Arg Ser Gly Pro Arg Leu Pro Gly Ala
20 25 30

Leu Leu Pro Ala Ala Ser Gly Pro Leu Gln Leu Arg Ala Leu Arg Gln
35 40 45

Gln Asp Leu Pro Ser Ala Leu Pro Gly Val Gly Gln Val Leu Gly Pro
50 55 60

Gly Arg Gly Ala His Leu Leu Leu His Trp Glu Arg Gly Arg Arg Val
65 70 75 80

Gly Leu Arg Gln Gln Leu Gly Leu Arg Arg Gly Leu Ala Ala Glu Arg
85 90 95

Gly Ala Leu Leu Val Phe Ala Glu His Arg Tyr Tyr Gly Lys Ser Leu
100 105 110

Pro Phe Gly Ala Gln Ser Thr Gln Arg Gly His Thr Glu Leu Leu Thr
115 120 125

Val Glu Gln Ala Leu Ala Asp Phe Ala Glu Leu Leu Arg Ala Leu Arg
130 135 140

Arg Asp Leu Gly Ala Gln Asp Ala Pro Ala Ile Ala Phe Gly Gly Ser
145 150 155 160

Tyr Gly Gly Met Leu Ser Ala Tyr Leu Arg Met Lys Tyr Pro His Leu
165 170 175

251
 Val Ala Gly Ala Leu Ala Ala Ser Ala Pro Val Leu Ser Val Ala Gly
 180 185 190
 Leu Gly Asp Ser Asn Gln Phe Phe Arg Asp Val Thr Ala Asp Phe Glu
 195 200 205
 Gly Gln Ser Pro Lys Cys Thr Gln Gly Val Arg Glu Ala Phe Arg Gln
 210 215 220
 Ile Lys Asp Leu Phe Leu Gln Gly Ala Tyr Asp Thr Val Arg Trp Glu
 225 230 235 240
 Phe Gly Thr Cys Gln Pro Leu Ser Asp Glu Lys Asp Leu Thr Gln Leu
 245 250 255
 Phe Met Phe Ala Arg Asn Ala Phe Thr Val Leu Ala Met Met Asp Tyr
 260 265 270
 Pro Tyr Pro Thr Asp Phe Leu Gly Pro Leu Pro Ala Asn Pro Val Lys
 275 280 285
 Val Gly Cys Asp Arg Leu Leu Ser Glu Ala Gln Arg Ile Thr Gly Leu
 290 295 300
 Arg Ala Leu Ala Gly Leu Val Tyr Asn Ala Ser Gly Ser Glu His Cys
 305 310 315 320
 Tyr Asp Ile Tyr Arg Leu Tyr His Ser Cys Ala Asp Pro Thr Gly Cys
 325 330 335
 Gly Thr Gly Pro Asp Ala Arg Ala Trp Asp Tyr Gln Ala Cys Thr Glu
 340 345 350
 Ile Asn Leu Thr Phe Ala Ser Asn Asn Val Thr Asp Met Phe Pro Asp
 355 360 365
 Leu Pro Phe Thr Asp Glu Leu Arg Gln Arg Tyr Cys Leu Asp Thr Trp
 370 375 380
 Gly Val Trp Pro Arg Pro Asp Trp Leu Leu Thr Ser Phe Trp Gly Gly
 385 390 395 400
 Asp Leu Arg Ala Ala Ser Asn Ile Ile Phe Ser Asn Gly Asn Leu Asp
 405 410 415
 Pro Trp Ala Gly Gly Gly Ile Arg Arg Asn Leu Ser Ala Ser Val Ile
 420 425 430
 Ala Val Thr Ile Gln Gly Gly Ala His His Leu Asp Leu Arg Ala Ser
 435 440 445
 His Pro Glu Asp Pro Ala Ser Val Val Glu Ala Arg Lys Leu Glu Ala
 450 455 460
 Thr Ile Ile Gly Glu Trp Val Lys Ala Ala Arg Arg Glu Gln Gln Pro
 465 470 475 480
 Ala Leu Arg Gly Gly Pro Arg Leu Ser Leu
 485 490

252

<210> 489
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 489
 Cys Ser Val Phe Pro Pro Ser Leu Trp Phe Tyr Leu Pro Leu Val Phe
 1 5 10 15
 Asp Asp Gly Asp Val Gln
 20

<210> 490
 <211> 122
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (46)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (113)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 490
 Gly Val Ser Leu Pro Leu Leu Gly Asp Ala Ser Gln Leu Gly Tyr Leu
 1 5 10 15
 Gly Val Arg Asp Ala Leu Glu Glu Ala Leu Cys Leu Phe Ser Asp Val
 20 25 30
 Gln Leu Cys Ala Gly Arg Thr Ser Ala Leu Phe Lys Ala Xaa Arg Gln
 35 40 45
 Gly Arg Leu Ser Leu Gln Arg Ile Leu Leu Pro Phe Val Trp Leu Cys
 50 55 60
 Pro Ala Pro Gln Arg Trp Ser Leu Gln Arg Gln Ala Gly Leu Leu Glu
 65 70 75 80
 Leu Arg Trp Ala Pro Pro Ser Ser Ser Phe Leu Ala Ala Leu Phe Thr
 85 90 95
 Pro Ser Ser Leu Gly Asn Gly Gly Arg Pro Ser Pro Ser Leu Thr Ala
 100 105 110
 Xaa Leu Gln Phe Asp Leu Arg Leu Leu Cys
 115 120

<210> 491
 <211> 74
 <212> PRT
 <213> Homo sapiens

<220>

253

<221> SITE

<222> (62)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (74)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 491

Val Cys Arg Gly Phe Cys Cys Leu Leu Phe Gly Cys Ala Leu Pro Pro
 1 5 10 15

Arg Gly Gly Val Tyr Arg Gly Arg Gln Ala Ser Leu Asn Cys Gly Gly
 20 25 30

Leu His Arg Val Arg Val Ser Trp Pro Leu Cys Leu Pro Pro Gln Ala
 35 40 45

Ser Ala Met Val Gly Ala Pro Pro Pro Ala Ser Leu Pro Xaa Cys Ser
 50 55 60

Leu Ile Ser Asp Cys Cys Ala Ser Asn Xaa
 65 70

<210> 492

<211> 34

<212> PRT

<213> Homo sapiens

<400> 492

Met Ser His Lys His Met Arg Arg Ser Ala Thr Ser Tyr Ile Ile Arg
 1 5 10 15

Glu Arg Gln Ile Lys Ile Ile Val Arg Tyr His Tyr Thr Pro Ile Met
 20 25 30

Thr Thr

<210> 493

<211> 16

<212> PRT

<213> Homo sapiens

<400> 493

Ile Arg Glu Arg Gln Ile Lys Ile Ile Val Arg Tyr His Tyr Thr Pro
 1 5 10 15

<210> 494

<211> 13

<212> PRT

<213> Homo sapiens

<400> 494

Lys Lys Thr Cys Thr Met Phe Ile Ala Thr Leu Phe Thr

1 5 254
 10
 <210> 495
 <211> 13
 <212> PRT
 <213> Homo sapiens
 <400> 495
 Glu Lys Ile Phe Ala Lys His Leu Ser Val Lys Gly Leu
 1 5 10
 <210> 496
 <211> 85
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (21)
 <223> Xaa equals any of the naturally occurring L-amino acids
 <220>
 <221> SITE
 <222> (39)
 <223> Xaa equals any of the naturally occurring L-amino acids
 <400> 496
 Ser Val Ala Ser Val Phe Ile Pro Leu Lys Val Ser Val Thr Lys Gln
 1 5 10 15
 Phe Ile Phe Phe Xaa Phe Phe Phe Phe Leu Arg Arg Ser Leu Ala Pro
 20 25 30
 Ala Trp Val Ala Glu Arg Xaa Thr Ser Gln Glu Thr Lys Gln Asn Lys
 35 40 45
 Lys Thr Pro Gln Leu Arg Gly Lys Val Ala His Ala Cys Asp Pro Ile
 50 55 60
 Thr Leu Gly Gly Arg Arg Trp Glu Val Gly Glu Ser Leu Glu Ala Arg
 65 70 75 80
 Ser Pro Ser Xaa Xaa
 85
 <210> 497
 <211> 184
 <212> PRT
 <213> Homo sapiens
 <400> 497
 Tyr Met Cys Cys Pro Phe Val Leu Asp Lys Asp Gly Val Ser Ala Ala
 1 5 10 15
 Val Ile Ser Ala Glu Leu Ala Ser Phe Leu Ala Thr Lys Asn Leu Ser
 20 25 30
 Leu Ser Gln Gln Leu Lys Ala Ile Tyr Val Glu Tyr Gly Tyr His Ile
 35 40 45

255

Thr Lys Ala Ser Tyr Phe Ile Cys His Asp Gln Glu Thr Ile Lys Lys
 50 55 60
 Leu Phe Glu Asn Leu Arg Asn Tyr Asp Gly Lys Asn Asn Tyr Pro Lys
 65 70 75 80
 Ala Cys Gly Lys Phe Glu Ile Ser Ala Ile Arg Asp Leu Thr Thr Gly
 85 90 95
 Tyr Asp Asp Ser Gln Pro Asp Lys Lys Ala Val Leu Pro Thr Ser Lys
 100 105 110
 Ser Ser Gln Met Ile Thr Phe Thr Phe Ala Asn Gly Gly Val Ala Thr
 115 120 125
 Met Arg Thr Ser Gly Thr Glu Pro Lys Ile Lys Tyr Tyr Ala Glu Leu
 130 135 140
 Cys Ala Pro Pro Gly Asn Ser Asp Pro Glu Gln Leu Lys Lys Glu Leu
 145 150 155 160
 Asn Glu Leu Val Ser Ala Ile Glu Glu His Phe Phe Gln Pro Gln Lys
 165 170 175
 Tyr Asn Leu Gln Pro Lys Ala Asp
 180
 <210> 498
 <211> 199
 <212> PRT
 <213> Homo sapiens
 <400> 498
 Ala Arg Gly Lys Thr Val Leu Phe Ala Phe Glu Glu Ala Ile Gly Tyr
 1 5 10 15
 Met Cys Cys Pro Phe Val Leu Asp Lys Asp Gly Val Ser Ala Ala Val
 20 25 30
 Ile Ser Ala Glu Leu Ala Ser Phe Leu Ala Thr Lys Asn Leu Ser Leu
 35 40 45
 Ser Gln Gln Leu Lys Ala Ile Tyr Val Glu Tyr Gly Tyr His Ile Thr
 50 55 60
 Lys Ala Ser Tyr Phe Ile Cys His Asp Gln Glu Thr Ile Lys Lys Leu
 65 70 75 80
 Phe Glu Asn Leu Arg Asn Tyr Asp Gly Lys Asn Asn Tyr Pro Lys Ala
 85 90 95
 Cys Gly Lys Phe Glu Ile Ser Ala Ile Arg Asp Leu Thr Thr Gly Tyr
 100 105 110
 Asp Asp Ser Gln Pro Asp Lys Lys Ala Val Leu Pro Thr Ser Lys Ser
 115 120 125
 Ser Gln Met Ile Thr Phe Thr Phe Ala Asn Gly Gly Val Ala Thr Met

130 135 256 140
Arg Thr Ser Gly Thr Glu Pro Lys Ile Lys Tyr Tyr Ala Glu Leu Cys
145 150 155 160
Ala Pro Pro Gly Asn Ser Asp Pro Glu Gln Leu Lys Lys Glu Leu Asn
165 170 175
Glu Leu Val Ser Ala Ile Glu Glu His Phe Phe Gln Pro Gln Lys Tyr
180 185 190
Asn Leu Gln Pro Lys Ala Asp
195

<210> 499
<211> 18
<212> PRT
<213> Homo sapiens

<400> 499
Asp Lys Asp Gly Val Ser Ala Ala Val Ile Ser Ala Glu Leu Ala Ser
1 5 10 15

Phe Leu

<210> 500
<211> 13
<212> PRT
<213> Homo sapiens

<400> 500
Arg Asp Leu Thr Thr Gly Tyr Asp Asp Ser Gln Pro Asp
1 5 10

<210> 501
<211> 15
<212> PRT
<213> Homo sapiens

<400> 501
Lys Ala Val Leu Pro Thr Ser Lys Ser Ser Gln Met Ile Thr Phe
1 5 10 15

<210> 502
<211> 17
<212> PRT
<213> Homo sapiens

<400> 502
Thr Met Arg Thr Ser Gly Thr Glu Pro Lys Ile Lys Tyr Tyr Ala Glu
1 5 10 15

Leu

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>199</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>August 28, 1997</u>	Accession Number <u>209226</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
Europe In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application <u>Elnora Rivera</u> Authorized PCT Operations - IAPD Team 1 (703) 305-3672 (703) 305-3230 (FAX)	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
--	--

ATCC Deposit 209226

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit No. 209226

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>201</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 7, 1998</u>	Accession Number <u>209852</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
Europe In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer <u>Enora Rivera</u> <u>PCT Operations - IAPD Team 1</u> <u>(703) 305-3678 (703) 305-3230 (FAX)</u>	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
--	--

ATCC Deposit 209852

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit No. 209852

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>204</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>May 7, 1998</u>	Accession Number <u>209853</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
Europe In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer <u>Elnora Rivera</u> <u>PCT Operations - IAPD Team 1</u> <u>(703) 305-3670 (703) 305-3230 (FAX)</u>	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
---	--

ATCC Deposit 209853

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit No. 209853

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>200</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>March 13, 1997</u>	Accession Number <u>97958</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
Europe In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only
<input checked="" type="checkbox"/> This sheet was received with the international application <u>Enora Rivera</u> Authorized officer <u>PCT Operations - IAPD Team 1</u> <u>(703) 305-3678 (703) 305-3230 (FAX)</u>

For International Bureau use only
<input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

ATCC Deposit 97958

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit No. 97958

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>198</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>April 20, 1998</u>	Accession Number <u>209782</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
Europe In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="display: flex; align-items: center;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border-top: 1px solid black; padding-top: 5px;"> <u>Elnora Rivera</u> Authorized officer - IAPD Team 1 <u>(703) 305-3678 (703) 305-3230 (FAX)</u> </div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="display: flex; align-items: center;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border-top: 1px solid black; padding-top: 5px;"> Authorized officer </div>
---	---

ATCC Deposit 209782

Page 2

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

ATCC Deposit No. 209782

Page 3

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/13418

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12N 15/12, 15/63, 1/21, 5/00; C07K 7/00, 14/435

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 69.3, 70.1, 325, 243, 320.1; 530/300, 350, 399; 536/23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, EMBASE, WPIDS, BIOSIS

search terms: secreted protein, antigenic, antigen

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JACOBS et al. A genetic selection for isolating cDNAs encoding secreted proteins. Gene. 1997, Vol. 198, pages 289-296.	1-12, 14-16, 21
X	US 5,534,409 A (GRONER et al) 09 July 1996, columns 21-26, especially see SEQ ID NO:2.	1-3, 7-11, 14-16

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

02 SEPTEMBER 1999

Date of mailing of the international search report

29 OCT 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

CHRISTINE SAOUD

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/13418**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-12, 14-16 and 21 with regard to SEQ ID NO:11, 130

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

435/69.1, 69.3, 70.1, 325, 243, 320.1; 530/300, 350, 399; 536/23.1

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-12, 14-16 and 21, drawn to polynucleotides, polypeptides, and recombinant methods of production.

Group II, claim(s) 13, drawn to an antibody.

Group III, claim(s) 17, drawn to methods of treatment by administering the polypeptide.

Group IV, claim(s) 17, drawn to methods of treatment by administering the polynucleotides.

Group V, claim(s) 18, drawn to methods of diagnosing by detecting the polynucleotide.

Group VI, claim(s) 19, drawn to methods of diagnosing by detecting the polypeptide.

Group VII, claim(s) 20, drawn to methods of determining a binding partner.

Group VIII, claim(s) 22, drawn to methods of identifying an activity in an assay.

Group IX, claim(s) 23, drawn to a binding partner.

In addition to the 11 groups listed above, each group is further directed to 94 distinct embodiments corresponding to the 94 pairs of sequence identifiers for the 94 different polynucleotides and polypeptides encoded thereby. Each polynucleotide and encoded polypeptides lack unity of invention because they do not share the same special technical feature. A special technical feature means those features that define a contribution which each of the claimed inventions, considered as a whole, makes over the prior art. The special technical feature of each polynucleotide is the specific nucleic acid sequence of the polynucleotide molecule. Unity of invention is found between the polynucleotide, the polypeptide and the recombinant methods of use of the polynucleotide to make the polypeptide because claims to these categories of invention all share the special technical feature of the polynucleotide.

The inventions listed as Groups II-IX do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the inventions of Groups II and IX do not share the special technical feature of Group I, which is the nucleic acid sequence of the polynucleotide. Groups III-VIII are directed to additional methods, however, PCT Article 17(3)(a) does not provide for multiple products, processes of manufacture or uses which are claimed. Therefore, the first invention of the category first mentioned in the claims of the application and the first recited invention of each of the other categories related thereto is considered the main invention of the claims.